An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia

Voelkle, Manuel C ; Oud, Johan H L ; Davidov, Eldad ; Schmidt, Peter

Abstract: Panel studies, in which the same subjects are repeatedly observed at multiple time points, are among the most popular longitudinal designs in psychology. Meanwhile, there exists a wide range of different methods to analyze such data, with autoregressive and cross-lagged models being 2 of the most well known representatives. Unfortunately, in these models time is only considered implicitly, making it difficult to account for unequally spaced measurement occasions or to compare parameter estimates across studies that are based on different time intervals. Stochastic differential equations offer a solution to this problem by relating the discrete time model to its underlying model in continuous time. It is the goal of the present article to introduce this approach to a broader psychological audience. A step-by-step review of the relationship between discrete and continuous time modeling is provided, and we demonstrate how continuous time parameters can be obtained via structural equation modeling. An empirical example on the relationship between authoritarianism and anomia is used to illustrate the approach.

DOI: https://doi.org/10.1037/a0027543

Originally published at:
Voelkle, Manuel C; Oud, Johan H L; Davidov, Eldad; Schmidt, Peter (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. Psychological Methods, 17(2):176-192.
DOI: https://doi.org/10.1037/a0027543

An SEM Approach to Continuous Time Modeling of Panel Data: Relating Authoritarianism and Anomia

Manuel C. Voelkle[1], Johan H.L. Oud[2], Eldad Davidov[3], Peter Schmidt[4]

[1]Max Planck Institute for Human Development, Germany – [2]Radboud University Nijmegen, The Netherlands – [3]University of Zurich, Zurich, Switzerland – [4]Higher School of Economics (HSE), Moscow, Russia

This is a pre-copy-editing, author-produced PDF published in the journal **Psychological Methods** Vol 17(2), June 2012, 176-192 *following peer review. This article may not exactly replicate the final version published in the APA journal. It is not the copy of record. The definitive publisher-authenticated version is available online under* **doi:10.1037/a0027543**

Acknowledgments

We would like to thank Timo von Oertzen for his helpful comments and stimulating discussions and Lisa Trierweiler for the English proof of the manuscript.

Person to Contact

Correspondence concerning this article should be addressed to Manuel C. Voelkle, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. E-mail: voelkle@mpib-berlin.mpg.de.

Abstract

Panel studies, where the same subjects are repeatedly observed at multiple time points, are among the most popular longitudinal designs in psychology. Meanwhile, there exists a wide range of different methods to analyze such data, with autoregressive and cross-lagged models being two of the most well-known representatives. Unfortunately, in these models time is only considered implicitly, making it difficult to account for unequally spaced measurement occasions or to compare parameter estimates across studies that are based on different time intervals. Stochastic differential equations offer a solution to this problem by relating the discrete time model to its underlying model in continuous time. It is the goal of the present article to introduce this approach to a broader psychological audience. A step-by-step review of the relationship between discrete and continuous time modeling is provided and we demonstrate how continuous time parameters can be obtained via structural equation modeling (SEM). An empirical example on the relationship between authoritarianism and anomia is used to illustrate the approach.

Keywords: continuous time modeling; stochastic differential equations; panel design; autoregressive cross-lagged model;

Outline

An SEM Approach to Continuous Time Modeling of Panel Data: Relating Authoritarianism and Anomia

How humans develop, how societies change over time, and what factors affect these changes are fundamental research topics in psychology and the social sciences. However, while in the real world most of these changes develop continuously, they usually cannot be observed in a truly continuous manner. Rather, researchers are forced to use "snapshots" of developmental processes in order to learn something about the underlying continuous time process and factors that possibly affect it.

Panel designs, in which the same subjects are repeatedly observed across time, are typical examples of such "snapshots". During the last decades a number of different methods have been developed to analyze panel data, each associated with specific strengths and weaknesses. Roughly, two broad categories of methods can be distinguished: models in which time is considered *explicitly* and models, in which time is only considered *implicitly*. Hierarchical linear models (e.g., Raudenbush & Bryk, 2002) or latent growth curve models (e.g., Bollen & Curran, 2006; Duncan, Duncan, & Strycker, 2006) are typical examples of the former. In hierarchical linear models (HLM) time is explicitly entered as a predictor into the model equation, while in latent growth curve models (LGM) time is represented by the factor loadings. In contrast, autoregressive and cross-lagged models are typical examples of the latter, because here time is only considered implicitly by the *order* of the measurement occasions, but not the exact time points or time intervals between them (e.g., Finkel, 1995). As a consequence, it is difficult to compare autoregressive and cross-lagged parameters that are based on different time intervals. To illustrate the problems associated with models for longitudinal data that do not explicitly account for time, let us consider three examples of increasing complexity:

1. On estimating and comparing autoregressive parameters

Imagine a researcher (Researcher 1), who is interested in the stability of physical well-being in children. His long term goal may be to compare physical well-being in children of different nationalities. For this purpose he has just finished a one year long panel study on South African children. For reasons of simplicity, let us assume physical well-being represents a single factor measured via self-report on a monthly basis (including the baseline measure, the study thus consists of 13 measurement occasions with a monthly interval of Δt_i = 1 month for all time intervals $i = 1,\ldots, T = 12$)[1]. In order to estimate stability, he opts for a very parsimonious model, which belongs to the most widely used models for longitudinal data: the autoregressive model of order one (cf. Lütkepohl, 2005):

$$x_i = ax_{i-1} + w_i. \qquad (1)$$

Because we will come back to this model several times in the following, it is worth having a closer look at it. In this simple form, physical well-being (x) at any discrete measurement occasion (i) is a function of the previous measurement occasion (x_{i-1}) weighted by the autoregressive coefficient (a), and an error term (w_i). If there is perfect stability a equals one, whereas $a = 0$, if physical well-being at any measurement occasion (i) is completely independent of previous well-being. Suppose, our researcher observed a stability coefficient of $a = 0.64$ indicating a moderate degree of stability. Suppose further, another researcher (Researcher 2) did exactly the same study, but with $T = 6$ and an interval of two months (Δt_i=

[1] When discussing differences between discrete and continuous time analyses, it is important to be precise with the notation. In the present paper we will use t to indicate the *exact time point of an observation*. For example, $x(t = 2011)$ means that variable x was observed in the year 2011, or $y(t = 105\text{ms})$ could mean that y was observed 105ms after a stimulus onset. The unit of t, of course, depends on the object of research. In contrast, $i = 1,\ldots, T$, is an index denoting the *rank of an observation* in a series of observations. In addition, we assume an initial measurement occasion ($i = 0$), which is predetermined. Because a study with $T + 1$ measurement occasions has T intervals between them, we can use the same index $i = 1,\ldots, T$ to indicate the (rank of a) time interval. For example, $x_{i=5}$ means that x was observed at the sixth (including the baseline measure at $i = 0$) measurement occasion. Of course the two notations may also be combined, with t_i indicating that time point t represents the $(i + 1)$'th measurement occasion in this study. For example, $x(t_{i=5} = 2011)$ means that x was observed in the year 2011, which constitutes the sixth measurement occasion in this study. The same logic applies to the *intervals* between two adjacent time points: Δt_i represents the i'th interval of length $t_i - t_{i-1}$. Note, that the first interval is denoted Δt_1 (and not Δt_0), while the first measurement of x is denoted x_0.

2). Researcher 2 observed a stability coefficient of a = 0.42. How can we compare the two coefficients? Is stability higher in Study 1 with a = 0.64 and Δt_i= 1 month or in Study 2 with a = 0.42 and Δt_i= 2 months?

As this little example suggests, the autoregressive coefficient a (and error term w for that matter) depends on the length of the time interval Δt_i between x_i and x_{i-1}. This information, however, is missing in Equation 1, because the autoregressive model considers time only "implicitly", by accounting for the *order* of the measurement occasions, but not for the *length* of the time intervals between them.

The situation gets even more complicated if Researcher 1, decides to extend his study by another year, but at a lower sampling rate of only 4 additional measurement occasions with an interval of Δt_i= 3 months between them. How can he estimate the stability of physical well-being (a) in such a design (i.e., Δt_i = 1 month during the first year and Δt_i = 2 months during the second year)? Even worse, Researcher 1 could decide to continue the study by handing out four identical questionnaires on physical well-being, and ask the children to return them at any four different time points throughout the next year. In this case, intervals would not only differ across time, but also across individuals. How can we compute the stability (a) in such a design? Simply ignoring the issue of different time intervals by applying Equation 1 is certainly not a solution.

2. On interpreting cross-lagged effects

Let us assume the two researchers are not just interested in the stability of physical well-being, but also in the relationship between physical and social well-being. In particular, they want to know whether physical well-being affects social well-being, whether social well-being affects physical well-being, or whether there is a reciprocal effect between the two constructs. Because the nature of the research question precludes a randomized experiment, the researchers choose a cross-lagged panel design as illustrated in Figure 1. Cross-lagged panel designs are often employed to infer causal relationships between variables (e.g., Granger, 1969). The underlying idea is to compare the cross-lagged effect of one construct *P* (physical well-being) measured at time point *t* on another construct *S* (social well-being) measured at time point *t* + 1, to the cross-lagged effect of *S* measured at *t* on *P* measured at *t* + 1.[2] Because temporally later events cannot cause earlier events, panel designs are considered a more suitable way to test causal relations than cross-sectional designs (Finkel, 1995; Granger, 1969; Oud, 2007b; but see also Rogosa, 1980, for a critique on the use of cross-lagged correlations).

The respective statistical model in Equation 2 is the multivariate extension of the autoregressive model introduced in Equation 1:

$$\mathbf{x}(t_i) = \mathbf{A}(\Delta t_i)\,\mathbf{x}(t_i - \Delta t_i) + \mathbf{w}(\Delta t_i). \tag{2}$$

However, there are two important differences: First, being a multivariate model $\mathbf{x}(t_i)$ is no longer a single variable, but a $V \times 1$ vector of outcome variables, with V being the number of variables observed at each time point t_i. In our case, $V = 2$, for physical and social well-being. Likewise, $\mathbf{A}(\Delta t_i)$ is now a $V \times V$ matrix relating the outcome variables over time. It contains the autoregressive effects in the main diagonal and cross-lagged effects in the off-diagonals. Finally, $\mathbf{w}(\Delta t_i)$ is a $V \times 1$ vector of prediction errors, which are assumed to be uncorrelated over time. Second—and more importantly—in contrast to Equation 1, Equation 2 makes it explicit that the autoregressive and cross-lagged effects $\mathbf{A}(\Delta t_i)$ depend on the time interval Δt_i between $\mathbf{x}(t_i)$ and $\mathbf{x}(t_i - \Delta t_i)$. The same applies to the error term $\mathbf{w}(\Delta t_i)$. Furthermore, the notation $\mathbf{x}(t_i)$ points to the fact that although $\mathbf{x}$ may be observed at any time point *t*, in an empirical study the time points are always discrete, thus the subscript *i* is added. Likewise, the time interval Δt_i between two discrete time points must also be discrete with $\Delta t_i = t_i -$

[2] For reasons of simplicity, we only consider first order AR(1) autoregressive effects (cf. Lütkepohl, 2005).

t_{i-1}. Note that just like a in Equation 1, $\mathbf{A}(\Delta t_i)$ is a function of the time interval (Δt_i) but apart from that the underlying process is assumed to be constant over time. At this point it is important to note that in Equation 2 we just highlight the fact that the parameters ($\mathbf{A}(\Delta t_i)$ and $\mathbf{w}(\Delta t_i)$) are actually functions of the time intervals Δt_i. The problem is that this information is not being used in standard autoregressive and cross-lagged models, which simply ignore time intervals. That is, in discrete time analysis Equation 2 is written simply as

$$\mathbf{x}_i = \mathbf{A}\,\mathbf{x}_{i-1} + \mathbf{w}_i.$$

As illustrated in Figure 1, as long as the two researchers use different time intervals (Researcher 1: Δt_i = 1 month versus Researcher 2: Δt_i = 2 months), they arrive not only at different conclusions regarding the stability of physical (and social) well-being, but also with respect to the cross-lagged effects between the two constructs. For both constructs, Researcher 1 observed larger stability coefficients than Researcher 2. In contrast, Researcher 2 observed stronger cross-lagged effects, in particular with respect to the effect of social on physical well-being (a_{12} = 0.18 for Δt_i = 1 month versus a_{12} = 0.27 for Δt_i= 2 months). In other examples, the relative size of the cross-lagged effects may even reverse, suggesting a change in the "causal" direction of effects. Based on such divergent results, it is easy to imagine the fierce debate in the scientific community on the true nature of the relationship between two (or more) constructs.

Of course the alert reader has long realized that there is no point in comparing parameter estimates that are based on different time intervals. In this example, the underlying process that generated the results of the two studies is exactly the same. The example is also quite realistic—in fact, the parameters were produced by the same model that also underlies the empirical example on the relationship between authoritarianism and anomia presented later in this article. The observed differences in discrete time parameter estimates are solely due to the fact that the two researchers used different time intervals in their studies. In order to find this out, however, we need to derive parameter estimates that are independent of the time intervals used in any specific study. Before we demonstrate how this is done, let us consider a final—and more abstract—example.

3. On finding the generating process

Let us assume Researcher 1 has the extraordinary ability to time travel. After having completed his study on the relationship between physical and social well-being with a time interval of Δt_i= 1 month, he travels back in time and repeats the study with a time interval of Δt_i = 0.5 months. As we have seen above, he obtained $\mathbf{A}(\Delta t_i = 1) = \begin{pmatrix} 0.64 & 0.18 \\ 0.03 & 0.89 \end{pmatrix}$ in his first attempt. When redoing the study, after traveling back in time, he obtained $\mathbf{A}(\Delta t_i = 0.5) = \begin{pmatrix} 0.80 & 0.10 \\ 0.02 & 0.94 \end{pmatrix}$. Because he traveled back in time, except for the different time intervals, everything else (i.e., the true relationship between physical and social well-being) was exactly the same as in the first study. Having obtained these different parameter estimates, he gets curious how the parameters will change if he repeats the process with many different intervals. So he travels forth and back in time, for say 1000 times, and records the parameter estimates for different intervals between Δt_i = 0 months and Δt_i = 10 months. He then plots the parameter estimates against the different time intervals and obtains Figure 2A for the autoregressive effects and Figure 2B for the cross-lagged effects. Being surprised by the systematic nature of the resulting plots he starts to wonder whether his time travels were really necessary or whether he could have known this in advance. Put more generally, given that the generating process (i.e., the process by which the two constructs and their relationship evolves over time) does not change, he wonders whether a single study would have been sufficient to compute this process.

As the reader may already suspect, the conjecture of Researcher 1 is correct. From now on we will refer to the generating process as the *continuous time model* (Bergstrom, 1984, 1988). Continuous time models are models in which time is not entered explicitly as an explanatory variable, but only as an index variable (see footnote 1). In contrast to discrete time models, however, the index variable may take on any continuous set of values.[3] In the following we will demonstrate how to estimate a continuous time model based on a discrete time panel study (without time traveling). Once the continuous time model is known, we can solve all of the problems raised in the examples above.

However, before we start introducing the approach a word of caution is in place: Time traveling is not possible! In the real world, we cannot repeat the exact same study with different intervals and be sure that (apart from the intervals) nothing changes. Rather, we have to *assume* that there is one underlying continuous time model (i.e., a single generating process). If we are willing to make this assumption, we can compute how different constructs and the relations among them look like for any arbitrary time interval. If we are not willing to make this assumption, the only way to find out is to conduct a new study for every time interval one is interested in. The degree to which it is reasonable to make this assumption depends on the research question at hand and is ultimately up to the researcher to decide. If we are willing to assume that there is one generating process underlying the relationship between physical and social well-being, continuous time modeling allows us to compare the results of studies that have been conducted with different time intervals, permits the computation of effects in studies with varying time intervals (within the study), and may even allow us to inter- or extrapolate to other time intervals. If we are not willing to make this assumption, there is no point in comparing parameter estimates of different studies, or parameter estimates obtained at different time intervals within the same study, regardless

[3] One can also conceive of such a continuous time model as a dynamical system. At each point in time, the system has a specific configuration, but time itself never acts as an explanatory variable in the model. Instead the model itself is an explanatory model. Arguably, this is often a more realistic view of the world, as compared to models that include time explicitly as a predictor (e.g., hierarchical linear models or latent growth curve

which statistical method is being used. Put more generally, we have to assume that the object of research (to identify the generating process) is independent of the method (any specific study with specific time intervals).

The Present Article

Obviously, what is needed is an approach that brings time back into discrete time models where it is otherwise only considered implicitly. This can be done by using stochastic differential equations (SDE), and it is the goal of the present paper to introduce continuous time modeling based on stochastic differential equations to a broader psychological audience. Having said that, little of what will be presented in the following is actually new in the sense that the method did not exist before. Quite the contrary: Continuous time models based on differential equations have been around since Newton in the 17th century. Likewise, stochastic differential equations are well established in other disciplines, such as physics or engineering, and many mathematical problems associated with them have been resolved in the first half of the 20th century. However, with few rather technical exceptions (e.g., Oud & Jansen, 2000; Oud & Delsing, 2010; Singer, 1998), continuous time models are virtually absent in the psychological literature and even experienced quantitative psychologists routinely rely on discrete time models in the assumed absence of better alternatives. Accordingly, with the present article we (1) want to introduce psychologists to continuous time modeling and (2) demonstrate how continuous time models may overcome many of the problems of discrete time analyses. Furthermore, we (3) would like to facilitate the use of continuous time models and (4) enable readers to understand and critically evaluate the outcomes of continuous time analyses. To achieve the latter two goals, we present continuous time models within the general framework of structural equation modeling (SEM) which

---

model). For example, when using a latent growth curve model to study learning, we typically do not assume that time "causes" learning, although we mathematically model it that way.

most readers are familiar with and supplement the article by computer code that does the analysis. Finally, we provide an example of a continuous time model using an empirical data set.

We will proceed in the following order: First, we will quickly review the conventional autoregressive cross-lagged model for discrete measurement. Second, the underlying continuous time model will be introduced in a stepwise fashion. Third, a number of important extensions will be introduced, including the continuous time stochastic error process and continuous time intercepts. Fourth, after a short technical summary of the relationship between the continuous and discrete time model, which aims at the mathematically more advanced audience, the model will be translated into the commonly known SEM framework. Fifth, an example on the relationship between authoritarianism and anomia will be provided to illustrate the approach. Finally, after considering further extensions and suggesting additional reading, we will conclude with a discussion of the advantages and limitations of the method.

The Discrete Time Autoregressive Cross-Lagged Model

The discrete time autoregressive model, with or without cross-lagged effects, is one of the most often used methods for the analysis of change in psychology (Hertzog & Nesselroade, 2003; Jöreskog, 1979; McArdle, 2009). The basic model has already been introduced in Equation 2:

$$\mathbf{x}(t_i) = \mathbf{A}(\Delta t_i)\,\mathbf{x}(t_i - \Delta t_i) + \mathbf{w}(\Delta t_i) \qquad \text{(2, repeated)}$$

As defined before, $\mathbf{x}(t_i)$ is a $V \times 1$ vector of outcome variables, with V being the number of variables observed at each time point. The $V \times V$ matrix $\mathbf{A}(\Delta t_i)$ relates the outcome variables over time. It is a function of the time interval (Δt_i), but apart from that assumed to be time invariant (but see Oud & Jansen, 2000). The subscript i indicates that although time itself is continuous (t), observations are necessarily taken at discrete time points (t_i). For the moment let us further assume that all variables are in deviation form, so that there is no need for an intercept term in Equation 2 (i.e., $E[\mathbf{x}(t_i)] = \mathbf{0}$, for all t_i and all subjects). This assumption will be relaxed later on. Finally, $\mathbf{w}(\Delta t_i)$ is a $V \times 1$ vector of error terms. Using SEM, $\mathbf{x}$ does not have to be directly observed but may be latent (cf. McArdle, 2009). From a substantive point of view, this typically requires measurement invariance over time, but also opens up some flexibility with respect to the measurement error (co)variance structure.

The Continuous Time Auto- and Cross-Effects Model

The continuous time model underlying Equation 2 will be introduced in a stepwise fashion. To facilitate understanding we will start with an *intuitive introduction* of the basic idea, before deriving the exact model and explicating the relationship between continuous and discrete time parameters. For reasons of readability and comprehension we will not discuss the stochastic error term until later on.

An Intuitive Approach to Continuous Time Modeling

Consider again our introductory example o*n estimating and comparing autoregressive parameters* in which two researchers investigated the stability of physical health using an autoregressive model. Because the time intervals differed, we were hesitant (and rightly so) to compare the resulting autoregressive effects directly, but wondered whether stability is higher in Study 1 with $a = 0.64$ and Δt_i= 1 month, or in Study 2 with $a = 0.42$ and Δt_i= 2 months? An intuitively appealing solution to this problem could be to compute the difference between $\mathbf{x}(t_i)$ and $\mathbf{x}(t_i - \Delta t_i)$ and divide this difference by the length of the time interval (Δt_i). This "normalizes" the change from one measurement occasion to the next with respect to the length of the time interval between them and gives us the rate of change: the so-called difference quotient. Predicting this normalized difference instead of $\mathbf{x}(t_i)$ results in the difference equation

$$\frac{\Delta \mathbf{x}(t_i)}{\Delta t_i} = \mathbf{A}_* \, \mathbf{x}(t_i - \Delta t_i) \tag{3}$$

with $\Delta\mathbf{x}(t_i) = \mathbf{x}(t_i) - \mathbf{x}(t_i - \Delta t_i)$. The original autoregressive matrix $\mathbf{A}(\Delta t_i)$ and $\mathbf{A}_*$ are related by

$$\mathbf{A}_* = \frac{\mathbf{A}(\Delta t_i) - \mathbf{I}}{\Delta t_i} \text{ or inversely } \mathbf{A}(\Delta t_i) = \mathbf{A}_* \Delta t_i + \mathbf{I}, \tag{4}$$

with $\mathbf{I}$ being an identity matrix. In contrast to the autoregressive cross-lagged matrix $\mathbf{A}(\Delta t_i)$, the new matrix $\mathbf{A}_*$ is rendered relatively independent of the time interval and can therefore be compared across studies with different observation intervals. Indeed, $\mathbf{A}_*$ is already a very crude approximation of the underlying continuous time model (i.e., the generating process). In order to better distinguish between continuous time and discrete time parameters, in the following we will speak of auto-effects and cross-effects in the continuous case, as compared to autoregressive and cross-lagged effects in the discrete case. The new matrix $\mathbf{A}_*$ contains the continuous time auto-effects in the main diagonal and cross-effects in the off-diagonals. Having computed the difference equation for both studies, it is now possible to compare the strength of the effects. The continuous time auto-effects for construct P in Study 1 are $\mathbf{A}_* = (0.64 - 1)/1 = -0.36$, whereas in Study 2 $\mathbf{A}_* = (0.42 - 1)/2 = -0.29$. Note that due to the subtraction of $\mathbf{I}$ in the numerator, autoregressive parameters between 0 and 1 become negative when translated into continuous time auto-effects. In our example we see that—other than suggested by the autoregressive effects (0.64 in Study 1 and 0.42 in Study 2)—the difference between the two auto-effects (–0.36 in Study 1 and –0.29 in Study 2) is much smaller, and the effect appears to be even weaker (i.e., more negative) in Study 1 as compared to Study 2. By simply ignoring the different time intervals, we would have come to exactly the opposite conclusion.

Although this brief example involves only autoregressive effects and no cross-lagged effects, the matrix notation in Equation 3 indicates that it generalizes to all possible autoregressive and cross-lagged models with any arbitrary number of variables and time points. The approach is not only intuitively appealing, but may *sometimes* be a convenient way to translate discrete time parameters (i.e., $\mathbf{A}(\Delta t_i)$) into continuous time parameters (i.e., $\mathbf{A}_*$) in order to compare effects across studies with different observation intervals. But most importantly, it is easy to implement. In a first step, any conventional program can be used to fit a standard autoregressive cross-lagged model, while in a second step the parameters can be transformed into continuous time parameters using Equation 4. This two-step procedure has been termed the *indirect approach* by Hamerle, Nagl, and Singer (1991).

Unfortunately, the intuitive approach is associated with at least two serious shortcomings so that despite the intuitive appeal, its use is generally discouraged. First, the relationship $\mathbf{A}(\Delta t_i) = \mathbf{A}_* \Delta t_i + \mathbf{I}$ is only a very crude approximation of the relationship between the true continuous time matrix ($\mathbf{A}$)—the so-called drift matrix, which will be introduced in the next paragraph—and the discrete time autoregressive matrix $\mathbf{A}(\Delta t_i)$. Second, the indirect approach requires the time intervals *within* a study to be of equal length, in order to enable equality constraints among parameters of different intervals. Consequently, the intuitive approach is at most an imprecise ad hoc method to compare autoregressive effects across studies with different time intervals between, but equal time intervals within. From a didactical perspective, however, it prepares the ground for introducing the exact relationship between continuous time and discrete time modeling, because in principle continuous time modeling follows the same logic as outlined above. Having introduced the exact relationship, however, we strongly discourage the use of the intuitive approach in practice.

An Exact Direct Approach to Continuous Time Modeling

By accounting for the length of the time interval (Δt_i) when estimating parameters ($\mathbf{A}_*$), the difference equation (Equation 3) brings time back into autoregressive cross-lagged panel models. The time intervals, however, remain discrete as indicated by subscript *i*. As a thought experiment, we could imagine what would happen if we make the time intervals in Equation 3 smaller and smaller (i.e., $\Delta t_i \rightarrow 0$). Mathematically speaking, this corresponds to taking the derivative of $\mathbf{x}(t)$ with respect to time. Note that—being a thought experiment—we can now refer to the exact time point (*t*) and are no longer restricted to the discrete time points actually observed in a study. Just as in Equation 3, the unknown derivative $\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t}$ can then be predicted by $\mathbf{x}(t)$, weighted by the so-called drift matrix $\mathbf{A}$:

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{A}\,\mathbf{x}(t). \tag{5}$$

Equation 5 is a (nonstochastic) differential equation because the derivative of $\mathbf{x}(t)$ is a function of $\mathbf{x}(t)$ itself. Fortunately, it is not a very complicated differential function, so that it is possible to quickly identify the only function $\mathbf{x}(t)$, which satisfies Equation 5, to be

$$\mathbf{x}(t) = \mathrm{e}^{\mathbf{A}\cdot(t-t_0)}\mathbf{x}(t_0), \tag{6}$$

with $\mathbf{x}(t_0)$ representing the vector of (exogenous) outcome variables at initial time point t_0. More precisely, Equation 6 is the only unique solution of Equation 5. Proof of this relationship is given in Appendix A.

At some point, however, the continuous time coefficients in drift matrix $\mathbf{A}$ have to be related to the discrete time autoregressive parameters (i.e., $\mathbf{A}(\Delta t_i)$ in Equation 2). This is done by setting Equation 6 equal to Equation 2 (see Appendix B). For starting value $\mathbf{x}(t_0)$ $= \mathbf{x}(t_i - \Delta t_i)$ and $(t - t_0) = \Delta t_i$, this allows us to express the *exact* relationship between discrete and continuous time as

$$\mathbf{A}(\Delta t_i) = \mathrm{e}^{\mathbf{A}\cdot\Delta t_i} \tag{7}$$

As before, $\mathbf{A}(\Delta t_i)$ contains all autoregressive and cross-lagged parameters of the discrete time model, while drift matrix $\mathbf{A}$ contains the corresponding auto- and cross-effects of the underlying continuous time model. As apparent from Equation 7, the relationship between the two is not linear, but follows a highly nonlinear matrix exponential function, which causes the sometimes rather paradoxical relationships between discrete and continuous time.

In contrast to the *intuitive approach*, Equation 7 gives the exact relationship between autoregressive cross-lagged matrix $\mathbf{A}(\Delta t_i)$ and continuous time drift matrix $\mathbf{A}$. The relationship of this *exact approach* to the *intuitive approach* becomes clear when expressing Equation 7 in power series expansion. The exponential function is commonly defined as $e^a = \sum_{k=0}^{\infty} \frac{a^k}{k!} = 1 + a + \frac{1}{2!}a^2 + \frac{1}{3!}a^3 + \cdots$. Thus, we can also express Equation 7 in matrix notation as

$$\mathbf{A}(\Delta t_i) = \mathrm{e}^{\mathbf{A}\cdot\Delta t_i} = \boxed{\mathbf{I} + \mathbf{A}\cdot\Delta t_i} + \frac{1}{2!}\mathbf{A}^2 \cdot {\Delta t_i}^2 + \frac{1}{3!}\mathbf{A}^3 \cdot {\Delta t_i}^3 \cdots. \tag{8}$$

Written in this form, it is obvious that $\mathbf{A}_*$ in the *intuitive approach* (Equation 4) accounts for just the first part of the entire power series (dashed part in Equation 8). This is the reason why the intuitive approach is only a very crude approximation of the *exact* approach just introduced. The logic underlying the intuitive approach, which does not need differential calculus, and the exact approach involving the matrix exponential function, is the same. Likewise, $\mathbf{A}_*$ and $\mathbf{A}$ can be interpreted analogously.

Having introduced the exact relationship between discrete and continuous time modeling, the question remains how to estimate the continuous time drift matrix ($\mathbf{A}$) based on discrete time intervals in a given panel study. As before, it is tempting to adopt the *indirect approach*, that is to estimate the discrete time parameters ($\mathbf{A}(\Delta t_i)$) in a first step, and solve

Equation 7 for **A** in a second step. However, while the linear approximation is no longer an issue with this indirect exact approach, the second problem of how to deal with parameter constraints for different time intervals remains (Hamerle, et al., 1991). Thus, an approach is needed which constrains the discrete time parameters to the underlying continuous time parameters directly during estimation. Because of the simultaneous estimation of discrete and continuous time parameters we speak of a *direct* approach in contrast to the two-step *indirect* method. Before having a closer look at the direct approach and how it can be used to translate discrete time models into continuous time, however, we must first consider the stochastic error term, which has been deliberately ignored so far.

A More Realistic View of the World: Introducing Errors

Consider again discrete time Equation 2. It has been shown that a simple but crude and problematic way to translate discrete time parameters into a continuous time framework is to compute the difference equation (Equation 3). In principle, the same has to be done with the error process. In discrete time, the error process is assumed to follow a random walk through time. As the name already suggests, random walk refers to a process where the value at each time point ($\mathbf{w}(t_i)$) is an additive function of the value of the immediately prior time point ($\mathbf{w}(t_i - \Delta t_i)$) and an additional component **e** (i.e., $\mathbf{w}(t_i) = \mathbf{w}(t_i - \Delta t_i) + \mathbf{e}$). If all **e** are drawn from a standard multivariate normal distribution with $N(\mathbf{0},\mathbf{I})$, then for $\Delta t_i = 1$ all successively nonoverlapping increments $\Delta\mathbf{w}(t_i) = \mathbf{w}(t_i) - \mathbf{w}(t_i - \Delta t_i)$ are independent with covariance matrix = $\Delta t_i \mathbf{I}$. Without changing the nature of the process, error variances other than 1 over $\Delta t_i = 1$, or correlated error terms, can be obtained by premultiplying $\Delta\mathbf{w}(t_i)$ with $\mathbf{G}$, where $\mathbf{G}$ is the Cholesky triangle of the desired covariance matrix $\mathbf{Q}$ (i.e., $\mathbf{Q} = \mathbf{G}\mathbf{G}'$). Thus, we could complement Equation 3 by adding the error term $\mathbf{G}\frac{\Delta\mathbf{w}(t_i)}{\Delta t_i}$. All problems associated with the intuitive (difference) approach would of course remain the same.

As demonstrated in the previous section, the exact drift matrix **A** is obtained by computing the derivative of $\mathbf{x}(t)$ with respect to time (see Equation 5). In the same manner the continuous time error process $\mathbf{W}(t)$ can be put in derivative form:

$$\mathbf{G}\,\frac{\mathrm{d}\mathbf{W}(t)}{\mathrm{d}t}. \tag{9}$$

Capital letter **W** is used to denote the continuous time error process and to distinguish it from the discrete time error term (lowercase letter **w**) in Equation 2. The continuous time error process $\mathbf{W}(t)$ is the limiting form of the discrete time random walk, better known as Wiener process or Brownian motion. The Wiener process has three important defining properties: First, it has independent increments with distribution $N(\mathbf{0}, \Delta t)$ over interval Δt. Second, its initial value $\mathbf{W}(t_0) = \mathbf{0}$, and third, $\mathbf{W}(t)$ is continuous, both with probability 1. As for the discrete time error process, variances larger or smaller than 1 are possible by premultiplication with the Cholesky triangle **G**. This also allows the specification of possible covariances among the prediction errors. Because the error covariance matrix **Q**—which is also referred to as the *diffusion matrix*—and **G** contain the same information, for any given **G**, **Q** is also known and vice versa ($\mathbf{Q} = \mathbf{G}\mathbf{G}'$).

Unfortunately, taking the derivative of the Wiener process $\mathbf{W}(t)$ is not as simple as taking the derivative of $\mathbf{x}(t)$. At this point we do not want to go into mathematical details, but roughly speaking the problem is that in discrete time it is easy to define a random walk as a process of adding independent increments, which are randomly drawn from $N(\mathbf{0},\mathbf{I})$. However, by making the increments smaller and smaller (i.e., by approximating the derivative in continuous time with $\Delta t \rightarrow 0$), we eventually end up with an infinite variance. One could think of it in terms of test theory, where making a test longer reduces its error variance, while making the test shorter increases its error variance. If we could make the test infinitesimally small (i.e., if a single item would not be the smallest unit), its error variance would go to infinity. This situation has given rise to a substantial amount of mathematical research. To avoid the derivative, the stochastic differential equation is often formulated in differential form by multiplying both sides by dt and is then interpreted as a stochastic integral equation. Even though the resulting integral still does not follow the normal rules of integration, the good news is that mathematics has solved this problem a long time ago, and it can be shown that for initial value $\mathbf{x}(t_0)$ and any time interval ($\Delta t = t - t_0$) between $\mathbf{x}(t_0)$ and $\mathbf{x}(t)$ the solution of

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{A}\,\mathbf{x}(t) + \mathbf{G}\frac{\mathrm{d}\mathbf{W}(t)}{\mathrm{d}t} \tag{10}$$

is

$$\mathbf{x}(t) = \mathrm{e}^{\mathbf{A}\cdot(t-t_0)}\mathbf{x}(t_0) + \int_{t_0}^{t} \mathrm{e}^{\mathbf{A}\cdot(t-s)}\mathbf{G}\mathrm{d}\mathbf{W}(s) \tag{11}$$

$$\text{with } cov\left[\int_{t_0}^{t} \mathrm{e}^{\mathbf{A}\cdot(t-s)}\mathbf{G}\mathrm{d}\mathbf{W}(s)\right] = \int_{t_0}^{t} \mathrm{e}^{\mathbf{A}\cdot(t-s)}\mathbf{Q}\mathrm{e}^{\mathbf{A}'\cdot(t-s)}\mathrm{d}s = irow\left\{\mathbf{A}_{\#}^{-1}\left[\mathrm{e}^{\mathbf{A}_{\#}\cdot(t-t_0)} - \mathbf{I}\right]row\mathbf{Q}\right\}$$

$$\text{for } \mathbf{Q} = \mathbf{GG}' \text{ and } \mathbf{A}_{\#} = \mathbf{A}\otimes\mathbf{I} + \mathbf{I}\otimes\mathbf{A}.$$

Note that the first part of Equation 11 corresponds to Equation 6 (see also the proof of the first part in Appendix A). The variable of integration (s) has been chosen in order not to confuse it with the upper limit of integration t. Via the operator *row* the elements of matrix $\mathbf{Q}$ are put row-wise into a column vector, while *irow* represents the inverse operation (i.e., putting the elements back into a matrix). $\otimes$ denotes the Kronecker product. Equations 10 and 11 are described in more detail in Appendix C, while we refer the reader to Arnold (1974, p. 128–134), Ruymgaart and Soong (1985, p. 80–99), Oud and Jansen (2000), or Singer (1990) for mathematical details. Because of the stochastic component of the error term, Equation 10 is called a stochastic differential equation.

Introducing Intercepts

The basic stochastic differential model introduced in Equation 10 can be extended in various ways. In particular, until now all variables were assumed to be in deviation form, thus the model did not permit nonzero mean trajectories. This assumption can be relaxed by adding a $V \times 1$ continuous time intercept vector $\mathbf{b}$ to Equation 10. As before, solving for initial value $\mathbf{x}(t) = \mathbf{x}(t_0)$, allows us to write the expected value of $\mathbf{x}(t)$ at any time point t as

$$E[\mathbf{x}(t)] = \mathrm{e}^{\mathbf{A}\cdot(t-t_0)} E[\mathbf{x}(t_0)] + \mathbf{A}^{-1}\left[\mathrm{e}^{\mathbf{A}\cdot(t-t_0)} - \mathbf{I}\right]\mathbf{b}. \quad (12)$$

Because in general the corresponding elements in drift matrix $\mathbf{A}$ become negative for autoregressive parameters between 0 and 1, $\mathrm{e}^{\mathbf{A}\cdot(t-t_0)}$ approaches zero for increasing time intervals. Thus $E[\mathbf{x}(t)]$ approaches $-\mathbf{A}^{-1}\mathbf{b}$ as $(t - t_0) \rightarrow \infty$. The final mean vector $(-\mathbf{A}^{-1}\mathbf{b})$, to which the process eventually converges to, represents the so-called (stable) equilibrium position. Just like in any ordinary regression analysis it is possible to account for different mean trajectories of different groups by replacing the $V \times 1$ vector $\mathbf{b}$ by the product of a $V \times R$ matrix $\mathbf{B}$ and an $R \times 1$ vector of R exogenous variables $\mathbf{u}$ (cf. Oud & Delsing, 2010). In principle, the variables in vector $\mathbf{u}$ may either be continuous or represent dummy variables (e.g., to allow different mean trajectories for men and women). As before, we assume that $\mathbf{b}$ and $\mathbf{Bu}$ do not vary across time (but see Oud & Jansen, 2000). Due to lack of space, however, group differences will not be considered any further in the present paper.

On the Relationship Between Continuous and Discrete Time: A Summary

So far, we have introduced the logic and rationale of continuous time modeling in a stepwise fashion by aiming at readers who are new to continuous time modeling. The present section integrates the previous parts in a compact—and mathematically explicit—form. Essentially, continuous time modeling can be summarized in five steps: First, the discrete time model is formulated as usual. Second, the derivative with respect to time is computed, resulting in a stochastic differential equation. Third, the stochastic differential equation of step two is solved for any arbitrary time interval. Fourth, discrete time parameters are constrained according to step three during the estimation process. Fifth, the model is estimated. For this we will formulate it as a structural equation model. In the following we will shortly summarize these steps:

Step 1. The complete discrete time model corresponds to Equation 2, augmented by the intercept vector $\mathbf{b}(\Delta t_i)$:

$$\mathbf{x}(t_i) = \mathbf{A}(\Delta t_i)\mathbf{x}(t_i - \Delta t_i) + \mathbf{b}(\Delta t_i) + \mathbf{w}(\Delta t_i). \quad (13)$$

Step 2. Taking the derivative with respect to time, the corresponding stochastic differential equation, which has been introduced in a stepwise fashion throughout the previous parts of this article, is

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{A}\,\mathbf{x}(t) + \mathbf{b} + \mathbf{G}\frac{\mathrm{d}\mathbf{W}(t)}{\mathrm{d}t}. \quad (14)$$

Step 3. For $\mathbf{x}(t) = \mathbf{x}(t_0)$ and any time interval ($\Delta t = t - t_0$) between $\mathbf{x}(t_0)$ and $\mathbf{x}(t)$, the solution of Equation 14 is

$$\mathbf{x}(t) = \mathrm{e}^{\mathbf{A}\cdot(t-t_0)}\mathbf{x}(t_0) + \mathbf{A}^{-1}\left[\mathrm{e}^{\mathbf{A}\cdot(t-t_0)} - \mathbf{I}\right]\mathbf{b} + \int_{t_0}^{t} \mathrm{e}^{\mathbf{A}\cdot(t-s)}\mathbf{G}\mathrm{d}\mathbf{W}(s) \quad (15)$$

with $cov\left[\int_{t_0}^{t} \mathrm{e}^{\mathbf{A}\cdot(t-s)}\mathbf{G}\mathrm{d}\mathbf{W}(s)\right] = \int_{t_0}^{t} \mathrm{e}^{\mathbf{A}\cdot(t-s)}\mathbf{Q}\mathrm{e}^{\mathbf{A}'\cdot(t-s)}\mathrm{d}s = irow\left\{\mathbf{A}_{\#}^{-1}\left[\mathrm{e}^{\mathbf{A}_{\#}\cdot(t-t_0)} - \mathbf{I}\right]row\mathbf{Q}\right\}$

for $\mathbf{Q} = \mathbf{G}\mathbf{G}'$ and $\mathbf{A}_{\#} = \mathbf{A}\otimes\mathbf{I} + \mathbf{I}\otimes\mathbf{A}$.

Step 4. For $\mathbf{x}(t_0) = \mathbf{x}(t_i - \Delta t_i)$ and $(t - t_0) = \Delta t_i$ Equation 13 can be set equal to Equation 15, which yields the relationships between continuous and discrete time parameters. Having identified these relationships, the discrete time model can be expressed as a function of the underlying continuous time parameters:

$$\mathbf{x}(t_i) = \mathrm{e}^{\mathbf{A}\cdot\Delta t_i}\,\mathbf{x}(t_i - \Delta t_i) + \mathbf{A}^{-1}[\mathrm{e}^{\mathbf{A}\cdot\Delta t_i} - \mathbf{I}]\mathbf{b} + \mathbf{w}(\Delta t_i) \quad (16)$$

with covariance matrix $\mathbf{Q}(\Delta t_i)$ as defined above.

Step 5. All that remains to be done is to estimate the parameters in Equation 16 and there are different ways to do so (Oud & Singer, 2008). One way is to use structural equation modeling. SEM is a well-established and convenient way to obtain maximum likelihood parameter estimates if it is possible to reformulate Equation 16 as a structural equation model and minimize the well-known function

$$F_{ML} = \log|\mathbf{\Sigma}| + \mathbf{tr}(\mathbf{S\Sigma^{-1}}) - \log|\mathbf{S}| - (V + 1) \tag{17}$$

with V denoting the number of observed variables, $\mathbf{S}$ being the observed, and $\mathbf{\Sigma}$ the model implied augmented moment matrix. The reformulation of Equation 16 as a structural equation model is demonstrated in the next paragraph.

Continuous Time Modeling in SEM

In SEM one commonly distinguishes between a measurement part as defined in Equation 18 and a structural part as shown in Equation 19 (e.g., Jöreskog, 1973; Bollen, 1989; Byrne, 1998).

$$\mathbf{y} = \mathbf{\Lambda\eta} + \boldsymbol{\varepsilon} \quad \text{with } cov(\boldsymbol{\varepsilon}) = \mathbf{\Theta} \tag{18}$$

$$\boldsymbol{\eta} = \mathbf{B\eta} + \boldsymbol{\zeta} \quad \text{with } cov(\boldsymbol{\zeta}) = \mathbf{\Psi} \tag{19}$$

In the measurement model, vector $\mathbf{y}$ contains the manifest (i.e., directly observed) variables, which are related to the latent factors in vector $\boldsymbol{\eta}$, weighted by the factor loading matrix $\mathbf{\Lambda}$, plus the corresponding measurement error vector $\boldsymbol{\varepsilon}$, with error covariance matrix $\mathbf{\Theta}$. In the structural model, the variables of interest $\boldsymbol{\eta}$ are related to each other via matrix $\mathbf{B}$. The prediction errors are contained in vector $\boldsymbol{\zeta}$ with covariance matrix $\mathbf{\Psi}$. For reasons of simplicity, we do not consider the measurement model in the present paper, so that we can ignore Equation 18 by setting $\boldsymbol{\eta} = \mathbf{y}$. The approach, however, generalizes readily to more complex models including latent variables. As usual, we assume that $\boldsymbol{\varepsilon}$ is uncorrelated with $\boldsymbol{\zeta}$

and $\boldsymbol{\eta}$ and that $E[\boldsymbol{\varepsilon}] = E[\boldsymbol{\zeta}] = \mathbf{0}$. For more detailed information we refer the reader to Bollen (1989; see p. 14 and 20 for model definition and standard assumptions). As mentioned above, measurement invariance is important when tracking latent constructs over time (cf. Meredith, 1993; Vandenberg, 2002). Once $\boldsymbol{\Lambda}$, $\boldsymbol{\Theta}$, $\mathbf{B}$, and $\boldsymbol{\Psi}$ have been defined, it is easy to derive the model implied covariance matrix (Bollen, 1989; p. 325) and estimate parameters by minimizing Equation 17. So all that needs to be done is to define $\boldsymbol{\eta}$, the matrix of regression coefficients $\mathbf{B}$, and the vector of error terms $\boldsymbol{\zeta}$, respectively, the covariance matrix $\boldsymbol{\Psi}$, in terms of Equation 16.

For $T + 1$ time points with V constructs at each occasion (each possibly measured by different indicators), vector $\boldsymbol{\eta}$ consists of $(T + 1) \cdot V$ distinct elements. However, if we want to permit nonzero mean trends, $\boldsymbol{\eta}$ must be extended by an additional element (the unit variable which has 1 for all sample units), so that

$$\boldsymbol{\eta}' = ([\mathbf{x}(t_0)]' \quad [\mathbf{x}(t_1)]' \quad [\mathbf{x}(t_i)]' \quad \cdots \quad [\mathbf{x}(t_{i=T})]' \quad 1)'$$

with all vectors within $\boldsymbol{\eta}$ as defined before. Accordingly, the vector of error terms is

$$\boldsymbol{\zeta}' = ([\mathbf{x}(t_0) - \boldsymbol{\mu}_{\mathbf{x}(t_0)}]' \quad [\mathbf{w}(\Delta t_1)]' \quad [\mathbf{w}(\Delta t_i)]' \quad \cdots \quad [\mathbf{w}(\Delta t_{i=T})]' \quad 1)',$$

with $\boldsymbol{\mu}_{\mathbf{x}(t_0)}$ being the $V \times 1$ mean vector of the V constructs observed at the first time point, and $\mathbf{w}(\Delta t_i)$ being a vector of discrete time error terms as defined above.

Constraining the discrete time parameters in Equation 2, to the underlying continuous time parameters according to Equation 16, the two matrices $\mathbf{B}$ and $\boldsymbol{\Psi}$ become

$$\mathbf{B} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \boldsymbol{\mu}_{\mathbf{x}(t_0)} \\ \mathrm{e}^{\mathbf{A}\cdot\Delta t_1} & \mathbf{0} & & \mathbf{0} & \mathbf{0} & \mathbf{A}^{-1}[\mathrm{e}^{\mathbf{A}\cdot\Delta t_1} - \mathbf{I}]\mathbf{b} \\ \mathbf{0} & \mathrm{e}^{\mathbf{A}\cdot\Delta t_i} & & \mathbf{0} & \mathbf{0} & \mathbf{A}^{-1}[\mathrm{e}^{\mathbf{A}\cdot\Delta t_i} - \mathbf{I}]\mathbf{b} \\ \vdots & & \ddots & & & \\ \mathbf{0} & \mathbf{0} & & \mathrm{e}^{\mathbf{A}\cdot\Delta t_{i=T}} & \mathbf{0} & \mathbf{A}^{-1}[\mathrm{e}^{\mathbf{A}\cdot\Delta t_{i=T}} - \mathbf{I}]\mathbf{b} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & 0 \end{pmatrix}$$

and

$$\mathbf{\Psi} = \begin{pmatrix} \mathbf{\Phi}[\mathbf{x}(t_0)] & & & & & \\ \mathbf{0} & \mathbf{Q}(\Delta t_1) & & & & \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}(\Delta t_i) & & & \\ \vdots & & & \ddots & & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & & \mathbf{Q}(\Delta t_{i=T}) & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & & \mathbf{0} & 1 \end{pmatrix}.$$

$\mathbf{\Phi}[\mathbf{x}(t_0)]$ is the covariance matrix of the (exogenous) constructs at the first occasion. We do not assume any prediction error at the first occasion.

Having introduced the relationship between discrete and continuous time parameters and having demonstrated how the necessary constraints can be formulated within the SEM framework, the next section provides an empirical example of the approach.

An Empirical Example: Relating Authoritarianism and Anomia

Over the past five decades, the theoretical concepts of authoritarianism and anomia have played an important role in sociology and social psychology. To date, most researchers (e.g., Alba, Schmidt, & Wasmer, 2004; Altemeyer, 1996; Lutterman & Middleton, 1970; Scheepers, Felling, & Peters, 1992; Stenner, 1997) agree that authoritarianism reflects a) an individual preference for submission under authorities (authoritarian submission), b) a strict orientation along the perceived conventions of the ingroup (authoritarian conventionalism), and c) aggressive stances toward outgroups (authoritarian aggression). Anomia has been defined by Srole (1956) as consisting of five subdimensions: a) political powerlessness, b) social powerlessness, c) generalized socioeconomic retrogression, d) normlessness and meaninglessness, and e) social isolation. The direction of the causal relation between the two constructs, however, is still controversial. First, it was hypothesized that anomia leads to authoritarianism (Merton, 1949; Srole, 1956), because it was assumed that individuals who feel normless and meaningless adopt authoritarian attitudes in order to regain orientation in an environment that is perceived as increasingly complex and irritating. This view, however, was challenged by an alternative explanation proposed by McClosky and Schaar (1965), who suggested that authoritarianism causes anomia. According to McClosky and Schaar (1965), certain personality characteristics as reflected by authoritarianism lead to anomia, because the narrow-mindedness of authoritarian people confines their opportunities for social interactions with others (e.g., Schlueter, Davidov, & Schmidt, 2007). Both positions share the view that the two constructs are rather stable over time, even though authoritarianism is regarded as somewhat more stable than anomia, because it represents a personality characteristic in the broadest sense. In contrast, anomia, which reflects an attitude of disorientation, is more susceptible to displaying changes in the relative position of individuals over time.

In our illustrative application we will take up the controversy and try to answer the question whether anomia leads to authoritarianism (supporting the view of Merton, 1949; Srole, 1956) or whether authoritarianism leads to anomia (supporting the view of McClosky & Schaar, 1965). Furthermore, we are interested in the stability of the two constructs over time. Although no claim is made with respect to *causal* relationships in a strict sense, which would require an experimental design, panel data offer a good (and oftentimes the only) opportunity to come close to the experimental ideal (Finkel, 1995).

Sample and Measurement Instruments

Data are taken from a recent panel study of the German general population aged 16 years and older without an immigration background (see Heitmeyer, 2004). Computer-assisted interviews were conducted at five points of measurement in 2002, 2003, 2004, 2006, and 2008. Note that the first three assessment waves were one year apart, while the last two measurement occasions took place after a two-year interval. No measurements were obtained in 2005 and 2007.

Authoritarianism was measured by four items which were selected from an authoritarianism scale used in previous German studies (Schmidt, Stephan, & Herrmann, 1995). Items were presented on a 4-point rating scale providing response options from 1 (*agree totally*) to 4 (*do not agree at all*). The original response options were recoded, so that higher values indicate higher agreement. The item wording was “In order to preserve law and order, it is necessary to act harder against outsiders”, “One should punish criminal acts harder”, “One should be obedient and respectful to authorities”, and “One should be grateful to leaders who tell us what to do”. The average of the four items was used for all subsequent analyses. Anomia was measured by three 4-point rating scale items “Everything has become so much in disarray that one does not know where one actually stands“, “Matters have become so difficult these days that one does not know what is going on“, and “People were better off in the past“. Just like for authoritarianism, the average rating of all three items was computed.

Table 1 contains some descriptive information and an overview of the sample size across the five observation waves. A total of N = 2,722 persons participated in the study. Response rates were 43% in the second wave, 30% in the third wave, 48% in the fourth wave, and 21% in the last wave. Although the loss of participants is substantial, it is typical for longitudinal surveys like the present one. Full information maximum likelihood (FIML) was used to deal with missing values.

Discrete Time Autoregressive Cross-Lagged Model

To investigate the causal relationship between anomia and authoritarianism, a standard discrete time autoregressive cross-lagged model was fitted to the data. Figure 3 shows a pictorial representation of the model. As discussed before, this model does not account for the fact that the time intervals between observation waves differed and will, therefore, yield incorrect results. However, since we are interested in the differences between the (correct) continuous time model and the (incorrect) discrete time model, let us start with a discrete time model.

The model contains 14 parameters to be estimated: Two autoregressive and two cross-lagged effects, two prediction error variances, two intercepts, and one prediction error covariance. All parameters are constrained to equality over time. In addition, the two means, two variances, and the covariance of the latent measures at the first time point were freely estimated. All measurement errors were set to zero, thus reducing the analysis to manifest variables only. This was done for reasons of simplicity. As demonstrated in the theoretical part of this article, the model can be easily extended to any arbitrary number of indicators at each time point. Maximum likelihood parameter estimates are provided in Table 2. As expected, with an autoregressive coefficient of 0.869, authoritarianism is a very stable construct. Likewise, with an autoregressive coefficient of 0.589, the stability of anomia is somewhat lower, but still reasonably high. Both cross-lagged effects are significant, but the effect of authoritarianism on anomia (0.202) is much higher than the effect of anomia on authoritarianism (0.033). Being a standard structural equation model, all parameters can be interpreted as usual, so that we will not go into details at this point.

Instead, we take up the question raised at the beginning of the article on how to compare the observed effects to effects of other studies. Suppose another researcher would have conducted a similar study, but used different time intervals and thus obtained different parameter estimates. Are the differences solely due to the different time intervals or do parameters differ irrespective of the length of the time interval? With the present discrete time model we cannot answer this question. Furthermore, in the present analysis we simply ignored the fact that the first two time intervals were one year, whereas the last two time intervals were two years. Even if time intervals differ only slightly, in some cases this may lead to completely wrong results and conclusions when ignored, while in other situations different time intervals may have little effect on results. In order to find out, however, we have to move to continuous time analysis.[4]

Continuous Time Auto- and Cross-Effects Model with Unequal Intervals

As described above, the matrix exponential relationship $\mathbf{A}(\Delta t_i) = e^{\mathbf{A}\cdot\Delta t_i}$ between the discrete time matrix $\mathbf{A}(\Delta t_i)$ and the continuous time matrix $\mathbf{A}$ lies at the heart of the exact approach (see Equation 7). To our knowledge, at present Mx and OpenMx are the only SEM programs that allow the use of such nonlinear constraints on matrices. Thus, while the results of the discrete time model in the previous section can be obtained by using any common SEM program,[5] we used OpenMx (Boker, et al., 2011) for estimating the continuous time parameters. OpenMx is an open source R-based (R Development Core Team, 2010) structural equation modeling program, which is freely available. All program scripts for the analyses in the present paper are available for download at [http://OMITTED FOR BLIND REVIEW].

Results of the continuous time auto- and cross-effects model with unequal intervals are given in Table 3. Most importantly, the drift matrix $\mathbf{A}$ is

$$\mathbf{A} = \begin{pmatrix} -0.447 & 0.232 \\ 0.043 & -0.117 \end{pmatrix}.$$

As noted above, this is the same drift matrix which has been used to construct the introductory example on the relationship between physical and social well-being. Having computed the parameters of the underlying continuous time model, it becomes possible to compute the corresponding discrete time parameters at any arbitrary point in time (see Equation 16). For example, computing the discrete time autoregressive and cross-lagged effects for time interval $\Delta t_i = 1$ we obtain

---

[4] As pointed out by an anonymous reviewer, an alternative approach to account for (few) unequal intervals is the use of phantom variables (McArdle, 2009; Rindskopf, 1984). This allows equality constraints on the discrete time parameters, even if some time intervals differ within a study. All other problems associated with discrete time analyses, however, remain.

[5] Parameter estimates and standard errors reported in Table 2 were identical for AMOS (Arbuckle, 1995-2009), M*plus* (L. K. Muthén & Muthén, 1998–2010), and OpenMx (Boker, et al., 2011).

$$\mathrm{e}^{\mathbf{A}\cdot\Delta t_i} = \mathrm{e}^{\begin{pmatrix}-0.447 & 0.232 \\ 0.043 & -0.117\end{pmatrix}\cdot 1} = \mathbf{A}(\Delta t_i = 1) = \begin{pmatrix}0.643 & 0.176 \\ 0.033 & 0.893\end{pmatrix}.$$

Comparing $\mathbf{A}(\Delta t_i = 1)$ to the autoregressive and cross-lagged parameters of the discrete time model in Table 2, we find that the correct autoregressive parameters are somewhat higher (anomia: 0.643 vs. 0.589; authoritarianism: 0.893 vs. 0.869) than the discrete time parameters. In contrast, for $\Delta t_i = 2$,

$$\mathrm{e}^{\mathbf{A}\cdot\Delta t_i} = \mathrm{e}^{\begin{pmatrix}-0.447 & 0.232 \\ 0.043 & -0.117\end{pmatrix}\cdot 2} = \mathbf{A}(\Delta t_i = 2) = \begin{pmatrix}0.419 & 0.271 \\ 0.050 & 0.804\end{pmatrix}$$

the effects are lower (anomia: 0.419 vs. 0.589; authoritarianism: 0.804 vs. 0.869). Obviously, by ignoring the length of the time intervals, the parameters obtained via a standard autoregressive cross-lagged analysis, are nonlinear mixtures of the parameters obtained for $\Delta t_i = 1$ and $\Delta t_i = 2$. Even though the difference in the stability of anomia (0.64 for $\Delta t_i = 1$ vs. 0.419 for $\Delta t_i = 2$) is already substantial, parameters may differ even more for more complex designs with larger differences in time intervals. In these situations, parameters of standard discrete time models can no longer be interpreted in any meaningful way.

Having obtained the continuous time parameters, we may now also inter- or extrapolate to any time interval of interest—provided that such inter- or extrapolation is meaningful on substantive grounds. Because the drift matrix is identical to the drift matrix of the introductory example on the relationship between physical and social well-being, the relationship between drift matrix $\mathbf{A}$ and the autoregressive cross-lagged parameters $\mathbf{A}(\Delta t_i)$ is depicted in Figure 2. The only difference is that time intervals are now in years rather than months. Authoritarianism is represented by construct *S* and anomia by construct *P*. Figure 2A shows the autoregressive coefficients of anomia and authoritarianism, Figure 2B the cross-lagged effects, and Figure 2C the expected values of authoritarianism and anomia. For $\Delta t_i =$ 0, the values in Figure 2C correspond to the descriptive means of the first measurement occasion in the discrete time model (cf. Equation 12). Probably most striking is the effect of the choice of time interval on the autoregressive effects: Based on the discrete time parameters, anomia and authoritarianism would both appear to be fairly stable constructs in a study with time intervals of one year. In contrast, in a study based on six year intervals, the stability of anomia would be expected to be close to zero ($a_{anan} = 0.06$), whereas the autoregressive effect of authoritarianism would still be substantial ($a_{auau} = 0.55$; see Figure 2A). Likewise, in a study with half-year intervals one would likely conclude that neither of the two constructs has a strong effect on the other, while a study with a time interval of four years lends strong support to the hypothesis that authoritarianism causes anomia (see Figure 2B). Without information on continuous time parameters one would be left with such contradictory results. Knowing the underlying drift matrix, however, it is readily apparent that authoritarianism is not only a more stable construct (–0.117 vs. –0.447) but has also a stronger effect on anomia (0.232) than the other way round (0.043). Thus, our results support the hypothesis of McClosky and Schaar (1965) that it is more likely that authoritarianism causes anomia than vice versa.

Extensions and Further Reading

Because the primary goal of this article is to introduce continuous time modeling based on stochastic differential equations to a broader psychological audience, we limited ourselves to the continuous time model in its basic form. In recent years, however, the basic approach has been extended in various ways and ongoing research promises further advancements. In this section we briefly want to mention some of these extensions and developments.

Obviously, all formulae provided in the present article generalize readily to multiple parallel processes with none, some, or all cross- and auto-effects being freely estimated. Also, multiple indicators may be used at each time point as long as measurement invariance can be guaranteed (Meredith, 1993). In addition, it is straightforward to include predictors and random subject effects (so-called traits; cf. Oud & Jansen, 2000). Using SEM to estimate continuous time parameters allows us to make use of the full flexibility of modern latent variable models (B. Muthén, 2002), including a range of different estimators, different link functions between indicators and latent constructs, or multiple group analyses to name just a few. The approach has also been extended to time-varying drift matrices (Oud & Jansen, 2000). Furthermore, the general idea of continuous time modeling is not limited to standard (vector) autoregressive and cross-lagged models as used in this paper, but applies to most longitudinal models in the social sciences that consider time only implicitly by accounting for the *order* of measurement occasions, but not for the *length* of the intervals between them. To some degree this is also true for hybrid models, such as the autoregressive latent trajectory model (Bollen & Curran, 2004; Curran & Bollen, 2001; Delsing & Oud, 2008) or growth curve models with time-varying covariates (Bollen & Curran, 2006). Likewise, continuous time models are not limited to panel data, but apply equally to time series data of single subjects (cf. Molenaar & Campbell, 2009). At present, most commonly used psychological methods for the analysis of (individual) time series simply ignore the length of the time intervals between observations. This is particularly true when using lagged (block-Toeplitz) covariance matrices to fit P-technique models (Cattell, Cattell, & Rhymer, 1947; Molenaar & Nesselroade, 2009), dynamic factor analytic models (Molenaar, 1985), or recently developed unified SE-models (Kim, Zhu, Chang, Bentler, & Ernst, 2007; Gates, Molenaar, Hillary, Ram, & Rovine, 2010). While in principle, all of these models can be extended to account for different time intervals, to our knowledge, this has not been done yet. However, given that parameter estimates are inherently bound to the length of the time intervals, future research should focus on extending continuous time modeling to these approaches as well.

Finally, it should be mentioned that various approximations have been developed to avoid the matrix exponential function in Equation 6 (cf. Bergstrom, 1988; Oud, 2007b; Oud

& Delsing, 2010). These are all based on different approximations of the power series expansion in Equation 8.[6] One advantage of the approximations is that many of them can be implemented in standard SEM packages like M*plus*.[7] Given that the OpenMx syntax provided along with this article is free of charge and offers the *exact* solution, this advantage seems negligible. However, in combination with a recently proposed *oversampling technique* (Singer, in press), the approximations can be an efficient way to avoid estimation problems associated with procedures based on the eigenvalue decomposition of the drift matrix **A**. This is particularly true when working with complex eigenvalues. In the present paper we limited ourselves to asymptotically stable, nonosciallating models, that is, models with negative and real-valued eigenvalues of **A**. By allowing complex eigenvalues, however, continuous time modeling can also be used to estimate (possibly coupled and/or damped) oscillating processes (e.g., Oud, 2007a; Oud & Folmer, in press; Singer, in press). Last but not least, we did not consider *individually* varying time intervals. The use of oversampling to estimate oscillating and nonoscillating continuous time models with individually varying time intervals is discussed by Voelkle and Oud (submitted).

Discussion and Conclusions

Despite the fact that most real-world phenomena change continuously over time, usually few discrete measurement occasions are available to infer the underlying process. For this purpose, a number of different methods have been developed with autoregressive and cross-lagged models being two of the most well-known representatives. Unfortunately, these methods consider time only implicitly, by accounting for the *order* of measurement occasions, but not for the *length* of the time intervals between them. As illustrated by three

[6] The *intuitive approach* used at the beginning of the paper to introduce the basic idea of continuous time modeling is one such approximation—albeit not a good one.
[7] We provide some example M*plus* code using an approximate approach at http://OMITTED FOR BLIND REVIEW].

short examples at the beginning of the article, this is highly problematic. First, parameters of standard autoregressive models cannot be compared across studies with different time intervals. Second, it is difficult to estimate and compare parameter estimates (e.g., stability coefficients) that are based on different time intervals *within* the same study. Third, in cross-lagged studies, the *relative* size of the lagged effect of one variable A on another variable B, and vice versa, is highly dependent on the time interval (see Figure 2B). In some cases the effects may even reverse, leaving the researcher with the paradoxical situation that for one time interval A "causes" B, while for another time interval B "causes" A. Finally, a discrete time model is inherently bound to the time intervals in a given study. It tells us little about the generating process that caused the data independent of the specific time intervals a researcher happens to have used.

Continuous time models on the basis of stochastic differential equations overcome these limitations. Although these models are known for several decades, they are virtually absent from the psychological literature. Accordingly, it was the purpose of the present paper to introduce psychologists to continuous time modeling by providing a step-by-step introduction to the approach. In short, the idea is to take the derivative of a continuous time process with respect to time. By solving the resulting differential equation, the relationship between discrete and continuous time parameters can be computed. Knowing this relationship it becomes possible to constrain the parameters of a discrete time model for any arbitrary Δt_i to the underlying continuous time parameters when estimating the model. That way one obtains both (discrete and continuous time) parameter sets directly during estimation. Once the continuous time parameters are known, we can easily solve all problems mentioned above.

Although there are different ways to formulate and estimate continuous time models (e.g., via filter techniques, cf. Oud & Singer, 2008), in this article we used SEM. With SEM we can not only capitalize on the full flexibility of general latent variable modeling (e.g., B. Muthén, 2002), but we also used an approach which is familiar to most psychologists. In particular the same assumptions and limitations (e.g., in terms of the number subjects, number of variables, or distributional properties) apply to the models discussed in the present article as to any other structural equation model. By minimizing Equation 17, we obtain maximum likelihood parameter estimates, as well as the likelihood of the data given the entire model (the –2 log(L) value returned by the syntax provided with this article corresponds to minus two times the log of the likelihood reported by most other SEM programs). Comparing nested models via the likelihood ratio statistic allows the user to conduct significance tests on any parameter, or combination of parameters, he/she is interested in, as well as the computation of overall goodness of fit indices. This topic has been extensively discussed in the literature (e.g., Bollen, 1989; Bollen & Long, 1993; Hu & Bentler, 1998, 1999; Hu, Bentler, & Kano, 1992; Marsh, 2004).

The probably biggest drawback of continuous time modeling is that—as compared to other methods for the analysis of change currently used by psychologists—the mathematics behind it may appear somewhat daunting. Granted, this is true to some degree, but in the present article we showed that the basic idea underlying continuous time modeling is actually quite simple. Furthermore, with the relevant and freely available computer code (downloadable at http://OMITTED FOR BLIND REVIEW) at hand, the user does not have to worry about the correct implementation of the most complicated Equations 14, 15, and 16, but may simply specify his/her standard discrete time SE-model and the software returns the continuous time parameter estimates. For the models discussed in the present paper, differences in computation times are also negligible. The continuous time model of our empirical example took about 9sec to be estimated on a standard PC. However, due to the complex (matrix exponential) parameter constraints, the optimization process is more susceptible to nonconvergence as compared to simpler structural equation models. As optimizers in SEM packages are being continuously improved, in the near future this may no longer be a problem. For the time being, however, finding a converging model by using the direct exact approach depends on good starting values. Should the user experience problems in finding a converging model, we recommend fitting a discrete time model first and applying the simple Equation 4 to derive starting values, which should usually suffice. However, we are currently extending the program to give the user a choice of different approximations to Equation 8, including oversampling (Singer, in press), with which starting values are likely to be less of a problem. The extended version will also allow the modeling of coupled (damped and undamped) oscillators, as well as individually varying time intervals (Voelkle & Oud, submitted).

To illustrate the approach, we provided an empirical example on the relationship between authoritarianism and anomia. In the example, two competing theories on the relationship between the two constructs have been compared. As expected, anomia and authoritarianism were both found to be fairly stable over time, with authoritarianism showing a slightly stronger (i.e., less negative) auto-effect than anomia (–0.117 vs. –0.447). In addition, we found that although there was a small but significant continuous time effect of anomia on authoritarianism (0.043), the effect of authoritarianism on anomia was much larger (0.232), lending support to the theory of McClosky and Schaar (1965).

Let us finish the article with a word of caution and a more general comment. First the word of caution: When should discrete time analysis be preferred over continuous time analysis? From a mathematical point of view, the short answer to this question is: never. The continuous time model contains exactly the same information as the discrete time model and more. It accounts not only for the order of measurement occasions but also for the time intervals between them. Thus, knowing the continuous time parameters, it is easy to reconstruct the discrete time parameters. The opposite is not true: Knowing the discrete time parameters, may tell us very little about the underlying continuous time model (i.e., the generating process). From an applied perspective, however, it may be that there is no point in *interpreting* continuous time parameters. This may be the case because the process actually develops in discrete time steps, and/or because the length of time intervals (within and across studies) does not vary and—being a constant—contains no information. In these situations, one may as well use a discrete time model. More importantly, the user must be careful in using continuous time parameters to inter- or extrapolate to discrete time points that have not been observed. No matter which statistical method is being used, this is always a dangerous thing to do. Sometimes, we have no other option, for example when we want to compare parameters that have been obtained in a study with $\Delta t_i = 1$ month, to parameters that have been observed in a study with $\Delta t_i = 2$ months, as in our first introductory example. Without inter- or extrapolating the findings one of the two studies to the time interval of the other, no comparisons can be made and no cumulative knowledge can be generated. In other situations it seems better to avoid such comparisons from the very beginning. For example, relating a study on emotional stability at a level of minutes, to a study on emotional stability over years, may not seem like a reasonable thing to do, even though—from a mathematical point of view—continuous time modeling would allow us to do so.

What does this all mean for applied quantitative research in psychology? Science progresses by cumulating evidence. For this purpose, it is crucial to be able to compare findings of studies that investigate the same phenomenon but use different time intervals. Likewise it must be possible to compare parameter estimates that were obtained at different time intervals within the same study. With conventional standard autoregressive and cross-lagged models—which belong to the most widely used longitudinal research methods in psychology—this is not possible. Continuous time modeling overcomes these limitations. We hope that the present introduction stimulates researchers to apply the approach to their own data and thus help to produce *cumulative* knowledge.

References

Alba, R., Schmidt, P., & Wasmer, M. (Eds.). (2004). *Germans or foreigners? Attitudes toward ethnic minorities in post-reunification Germany*. New York: Palgrave Macmillan.

Altemeyer, B. (1996). *The authoritarian specter*. Cambridge, MA: Harvard University Press.

Arbuckle, J. L. (1995-2009). Amos 18.0 user's guide. Chicago, IL: SPSS.

Arnold, L. (1974). *Stochastic differential equations*. New York: Wiley.

Bergstrom, A. R. (1984). Continuous time stochastic models and issues of aggregation over time. In Z. Griliches & M. D. Intriligator (Eds.), *Handbook of econometrics.* (Vol. 2). Amsterdam: Elsevier Science.

Bergstrom, A. R. (1988). The history of continuous-time econometric models. *Econometric Theory, 4*, 365–383.

Boker, S. M., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., et al. (2011). OpenMx: An Open Source Extended Structural Equation Modeling Framework. *Psychometrika, 76*(2), 306–317.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons, Inc.

Bollen, K. A., & Curran, P. J. (2004). Autoregressive latent trajectory (ALT) models: A synthesis of two traditions. *Sociological Methods & Research, 32*, 336–383. doi: 10.1177/0049124103260222

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. Hoboken, NJ: John Wiley.

Bollen, K. A., & Long, S. J. (Eds.). (1993). *Testing structural equation models*. Newbury Park, CA: Sage.

Byrne, B. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts applications and programming*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Cattell, R. B., Cattell, A. K. S., & Rhymer, R. M. (1947). P-technique demonstrated in determining psychophysical source traits in a normal individual. *Psychometrika, 12*, 267–288. doi: 10.1007/BF02288941

Curran, P. J., & Bollen, K. A. (2001). The best of both worlds: Combining autoregressive and latent curve models. In L. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 107–135). Washington, DC: American Psychological Association.

Delsing, M. J. M. H., & Oud, J. H. L. (2008). Analyzing reciprocal relationships by means of the continuous-time autoregressive latent trajectory model. *Statistica Neerlandica, 62*(1), 58–82. doi: 10.1111/j.1467-9574.2007.00386.x

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Finkel, S. E. (1995). *Causal analysis with panel data*. Thousand Oaks: Sage.

Gates, K. M., Molenaar, P. C. M., Hillary, F. G., Ram, N., & Rovine, M. J. (2010). Automatic search for fMRI connectivity mapping: An alternative to Granger causality testing using formal equivalences among SEM path modeling, VAR, and unified SEM. *NeuroImage, 50*, 1118–1125. doi: 10.1016/j.neuroimage.2009.12.117

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*, 424–438. doi: 10.2307/1912791

Hamerle, A., Nagl, W., & Singer, H. (1991). Problems with the estimation of stochastic differential equations using structural equation models. *Journal of Mathematical Sociology, 16*, 201–220. doi: 10.1080/0022250X.1991.9990088

Heitmeyer, W. (Ed.). (2004). *Deutsche Zustände. Folge 3 [Current state in Germany. Series 3]*. Frankfurt a. M.: Suhrkamp.

Hertzog, C., & Nesselroade, J. R. (2003). Assessing psychological change in adulthood: An overview of methodological issues. *Psychology and Aging, 18*, 639–657. doi: 10.1037/0882-7974.18.4.639

Hu, L.-t., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*(4), 424–453.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55.

Hu, L.-t., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin, 112*(2), 351–362.

Jöreskog, K. G. (1973). A general method for estimating a linear structural equation system. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 85–112). New York: Seminar Press.

Jöreskog, K. G. (1979). Statistical estimation of structural models in longitudinal development investigations. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 303–352). New York: Academic Press.

Kim, J., Zhu, W., Chang, L., Bentler, P. M., & Ernst, T. (2007). Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. *Human Brain Mapping, 28*, 85–93. doi: 10.1002/hbm.20259

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Berlin: Springer.

Lutterman, K. G., & Middleton, R. (1970). Authoritarianism, anomia, and prejudice. *Social Forces, 48*, 485–492. doi: 10.2307/2575572

Marsh, H. W. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff value for fit indexes and dangers in overgeneralizing Hu and Bentler´s (1999) findings. *Structural Equation Modeling, 11*(3), 320–341.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577–605. doi: 10.1146/annurev.psych.60.110707.163612

McClosky, H., & Schaar, J. H. (1965). Psychological dimensions of anomy. *American Sociological Review, 30*, 14–40. doi: 10.2307/2091771

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525–543.

Merton, R. K. (1949). Social structure and anomie: Revisions and extensions. In R. Anshen (Ed.), *The family* (pp. 226–257). New York: Harper Brothers.

Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika, 50*(2), 181–202. doi: 10.1007/BF02294246

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychology, 18*(2), 112–117. doi: 10.1111/j.1467-8721.2009.01619.x

Molenaar, P. C. M., & Nesselroade, J. R. (2009). The recoverability of P-technique factor analysis. *Multivariate Behavioral Research, 44*(1), 130–141. doi: 10.1080/00273170802620204

Muthén, B. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika, 29*(1), 81–117. doi: 10.2333/bhmk.29.81

Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus users's guide* (Sixth ed.). Los Angeles, CA: Muthén & Muthén.

Oud, J. H. L. (2007a). Comparison of four procedures to estimate the damped linear differential oscillator for panel data. In K. van Montfort, J. H. L. Oud & A. Satorra (Eds.), *Longitudinal models in the behavioral and related sciences* (pp. 19–39). Mahwah, NJ: Lawrence Erlbaum Associates.

Oud, J. H. L. (2007b). Continuous time modeling of reciprocal relationships in the cross-lagged panel design. In S. M. Boker & M. J. Wenger (Eds.), *Data analytic techniques for dynamical systems* (pp. 87–129): Mahwah, NJ: Lawrence Erlbaum Associates.

Oud, J. H. L., & Delsing, M. J. M. H. (2010). Continuous time modeling of panel data by means of SEM. In K. van Montfort, J. H. L. Oud & A. Satorra (Eds.), *Longitudinal research with latent variables* (pp. 201–244). New York: Springer.

Oud, J. H. L., & Folmer, H. (in press). Modeling oscillation: approximately or exactly? *Multivariate Behavioral Research*.

Oud, J. H. L., & Jansen, R. A. R. G. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika, 65*(2), 199–215. doi: 10.1007/BF02294374

Oud, J. H. L., & Singer, H. (2008). Continuous time modeling of panel data: SEM versus filter techniques. *Statistica Neerlandica, 62*(1), 4–28.

R Development Core Team. (2010). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Rindskopf, D. (1984). Using phantom and imaginary latent variables to parameterize constraints in linear structural models. *Psychometrika, 49*(1), 37–47. doi: 10.1007/bf02294204

Rogosa, D. R. (1980). A critique of cross-lagged correlation. *Psychological Bulletin, 88*(2), 245–258. doi: 10.1037/0033-2909.88.2.245

Ruymgaart, P. A., & Soong, T. T. (1985). *Mathematics of Kalman-Bucy filtering*. Berlin: Springer.

Scheepers, P., Felling, A., & Peters, J. (1992). Anomie, authoritarianism, and ethnocentrism: Update of a classic theme and an empirical test. *Politics and the Individual, 2*, 43–60.

Schlueter, E., Davidov, E., & Schmidt, P. (2007). Applying autoregressive cross-lagged and latent growth models to a three-wave panel study In K. van Montfort, J. H. L. Oud & A. Satorra (Eds.), *Longitudinal models in the behavioral and related sciences* (pp. 315–336). Mahwah, NJ: Lawrence Erlbaum Publishers.

Schmidt, P., Stephan, K., & Herrmann, A. (1995). Entwicklung einer Kurzskala zur Messung von Autoritarismus [The development of a brief scale for the measurement of authoritarianism]. In G. Lederer & P. Schmidt (Eds.), *Autoritarismus und Gesellschaft. Trendanalysen und vergleichende Jugenduntersuchungen 1945–1993 [Authoritarianism and society: Trend analyses and comparative youth studies 1945–1993]* (pp. 221–227). Opladen: Leske & Budrich.

Singer, H. (1990). *Parameterschätzung in zeitkontinuierlichen dynamischen Systemen [Parameter estimation in continuous time dynamic systems]*. Konstanz: Hartung-Gorre.

Singer, H. (1998). Continuous panel models with time dependent parameters. *Journal of Mathematical Sociology, 23*(2), 77–89. doi: 10.1080/0022250X.1998.9990214

Singer, H. (in press). SEM modeling with singular moment matrices part II: ML-estimation of sampled stochastic differential equations. *Journal of Mathematical Sociology*.
Srole, L. (1956). Social integration and certain corollaries: An exploratory study. *American Sociological Review, 21*, 709–716. doi: 10.2307/2088422
Stenner, K. (1997). *Societal threat and authoritarianism: Racism, intolerance, and punitiveness in America, 1960–1994*. Ann Arbor: UMI.
Vandenberg, R. J. (2002). Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods., 5*(2), 139–158. doi: 10.1177/1094428102005002001
Voelkle, M. C., & Oud, J. H. L. (submitted). Continuous time modeling with individually varying time intervals for oscillating and nonoscillating models.

Footnotes

Will be inserted at the very end

Table 1

Sample Size (N), Mean (M), and Standard Deviation (SD) for Anomia and Authoritarianism Measured in 2002, 2003, 2004, 2006, 2008 (no Assessments in 2005 and 2007).

	Anomia			**Authoritarianism**		
	N	*M*	*SD*	*N*	*M*	*SD*
2002	2,721	2.50	0.80	2,722	2.84	0.68
2003	1,175	2.70	0.82	1,176	2.85	0.67
2004	826	2.81	0.78	826	2.83	0.67
2006	1,024	2.63	0.80	1,298	2.92	0.81
2008	560	2.48	0.80	1,047	2.70	0.82

Table 2

Parameter Estimates of the Discrete Time Autoregressive Cross-Lagged Model of Anomia and Authoritarianism across Four Time Intervals.

	Parameter	**Estimate**	**Standard Error**
Autoregressive effects	a_{anan}	0.589**	0.014
	a_{auau}	0.869**	0.009
Cross-lagged effects	a_{anau}	0.202**	0.015
	a_{auan}	0.033**	0.009
Latent intercepts	b_{an}	0.532**	0.041
	b_{au}	0.289**	0.026
Residuals	$var(w_{an})$	0.346**	0.008
	$var(w_{au})$	0.174**	0.004
	$cov(w_{auan})$	0.026**	0.004
Initial measurement occasion (means / covariances)	$M(an_{t0})$	2.503**	0.015
	$M(au_{t0})$	2.843**	0.013
	$var(an_{t0})$	0.633**	0.017
	$var(au_{t0})$	0.458**	0.012
	$cov(an_{t0}, au_{t0})$	0.245**	0.011
–2 log(L) (FIML)	23,073.60		

Note. **p < .01; *an* = anomia; *au* = authoritarianism; *anau* = regression of anomia on authoritarianism; *auan* = regression of authoritarianism on anomia.

Table 3

Parameter Estimates of the Continuous Time Auto- and Cross-Effects Model for Unequal

($\Delta t_1 = \Delta t_2 = 1;\ \Delta t_3 = \Delta t_4 = 2$*) Intervals.*

			Continuous Time Parameter Estimates	Standard Error
Drift matrix (**A**)	*Auto-effects*	a_{anan}	–0.447**	0.020
		a_{auau}	–0.117**	0.009
	Cross-effects	a_{anau}	0.232**	0.018
		a_{auan}	0.043**	0.010
Continuous time intercepts (**b**)	b_{an}		0.536**	0.046
	b_{au}		0.220**	0.022
Diffusion matrix (**Q**)	q_{anan}		0.473**	0.016
	q_{auau}		0.154**	0.004
	$q_{anau} = q_{auan}$		–0.005 n.s.	0.005
Initial measurement occasion (means / covariances)	$M(an_{t0})$		2.503**	0.015
	$M(au_{t0})$		2.843**	0.013
	$var(an_{t0})$		0.633**	0.017
	$var(au_{t0})$		0.458**	0.012
	$cov(an_{t0}, au_{t0})$		0.245**	0.011
Model fit (FIML)	–2 log(L)		23,415.93	
	df		13361	

Note. $^{**}p < .01$; n.s. = not significant; *an* = anomia; *au* = authoritarianism; *anau* = effect of authoritarianism on anomia; *auan* = effect of anomia on authoritarianism; FIML: Full information/raw data maximum likelihood estimation (FIML); See text for details.

Figure Captions

Figure 1. Autoregressive and cross-lagged parameter estimates of two studies on the relationship between two constructs across $T = 12$ versus $T = 6$ intervals. All parameter estimates were constrained to equality over time, and time intervals are assumed to be of equal length within each study ($\Delta t_1,\ldots, \Delta t_{12} = 1$ month in Study 1 and $\Delta t_1,\ldots, \Delta t_6 = 2$ months in Study 2).

Figure 2. A) Autoregressive parameters as a function of the time interval between observations. B) Cross-lagged parameters as a function of the time interval between observations. For a time interval of $\Delta t_{i\,=\,1,\,\ldots,\,T} = 1$, parameter estimates correspond to the discrete time effects observed by Researcher 1, while parameter estimates correspond to the discrete time effects observed by Researcher 2 for $\Delta t_{i\,=\,1,\ldots,\,T} = 2$. C) Expected values of authoritarianism (construct *S*) and anomia (construct *P*) as a function of time.

Figure 3. Bivariate discrete time autoregressive cross-lagged model of authoritarianism (*au*) and anomia (*an*). Squares represent observed (manifest) variables, circles/ellipses represent latent variables. For reasons of simplicity, all measurement errors were fixed to zero in the present example (dashed circles) making it a model with only manifest variables. However, as the figure illustrates, the model can be easily extended to latent variables. The triangle to the right represents the constant 1. Its path coefficients (one-headed arrows) represent the means or intercepts of the variables in question. Path coefficients associated with dashed lines are all fixed to one, while path coefficients associated with solid lines are freely estimated but are usually constrained to other parameters as described in the text. Double-headed arrows indicate covariances.

Figure 1

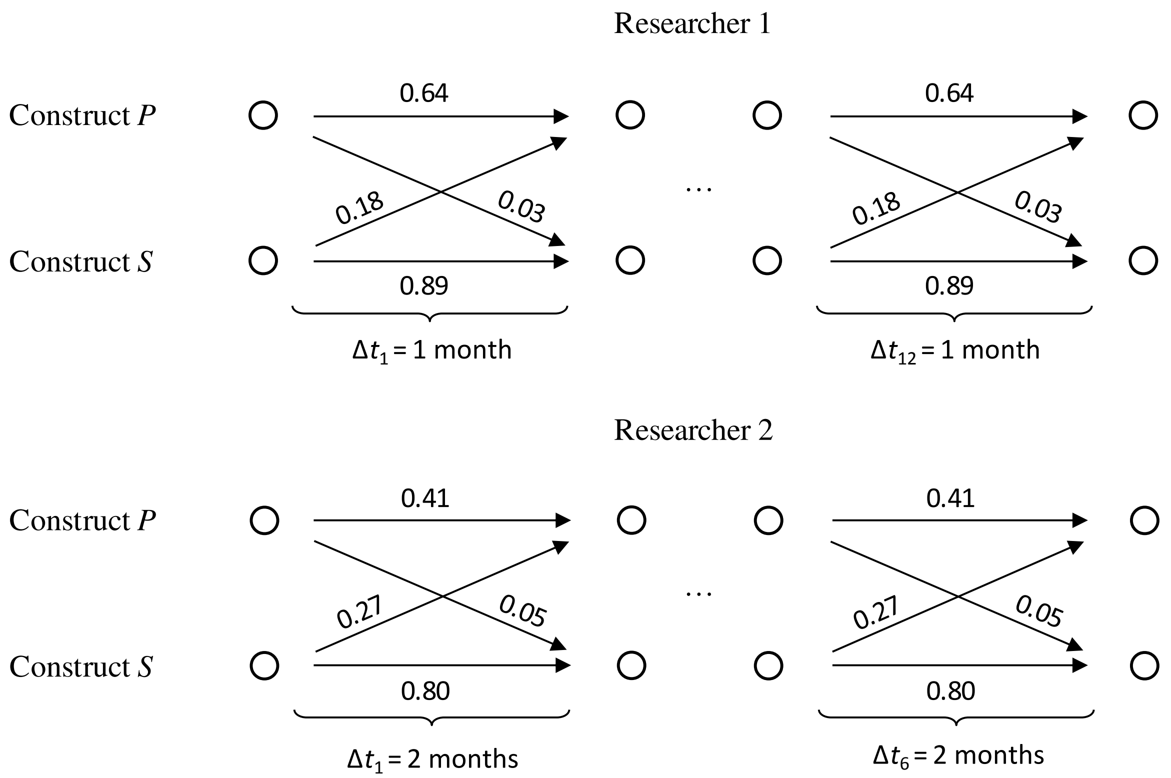
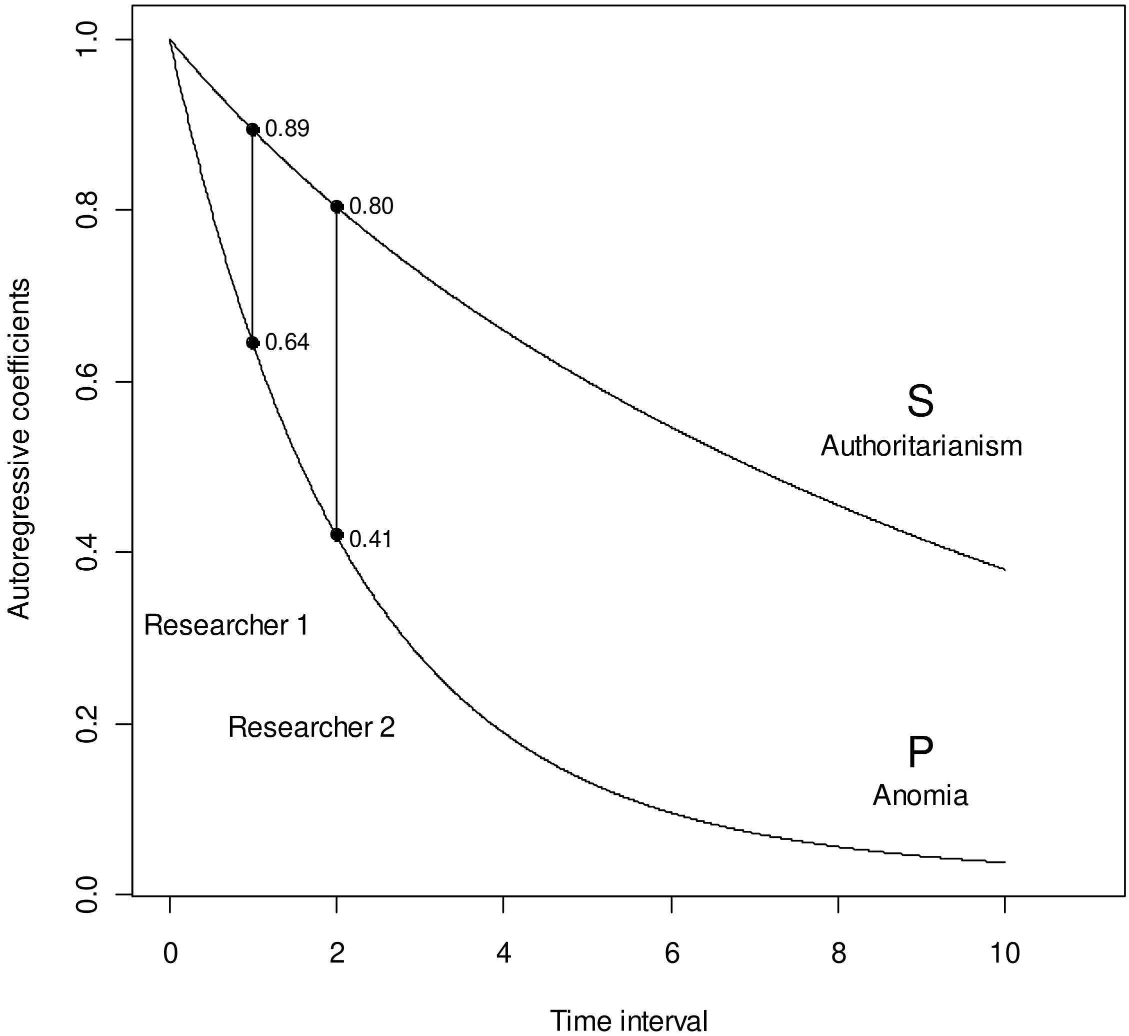
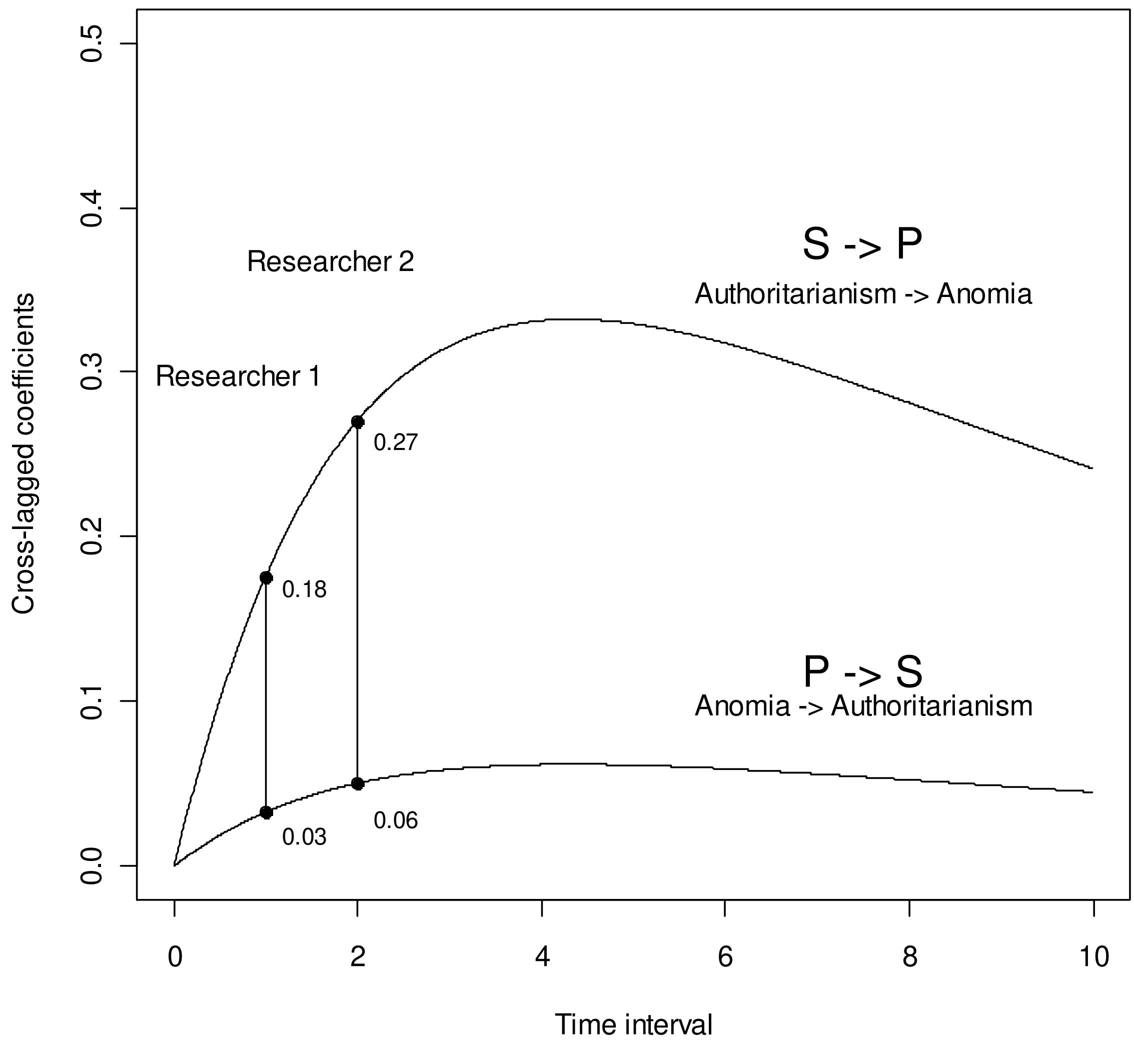

Researcher 1
Construct P
Construct S
0.64
0.18
0.03
0.89
Δt1 = 1 month
0.64
0.18
0.03
0.89
Δt12 = 1 month
Researcher 2
Construct P
Construct S
0.41
0.27
0.05
0.80
Δt1 = 2 months
0.41
0.27
0.05
0.80
Δt6 = 2 months

Figure 2A)



1.0
0.8
0.6
0.4
0.2
0.0
Autoregressive coefficients
0.89
0.80
0.64
0.41
S
Authoritarianism
Researcher 1
Researcher 2
P
Anomia
0
2
4
6
8
10
Time interval

Figure 2B)



S -> P
Authoritarianism -> Anomia
P -> S
Anomia -> Authoritarianism
Researcher 1
Researcher 2
0.27
0.18
0.03
0.06
Cross-lagged coefficients
0.0
0.1
0.2
0.3
0.4
0.5
0
2
4
6
8
10
Time interval

Figure 2C)

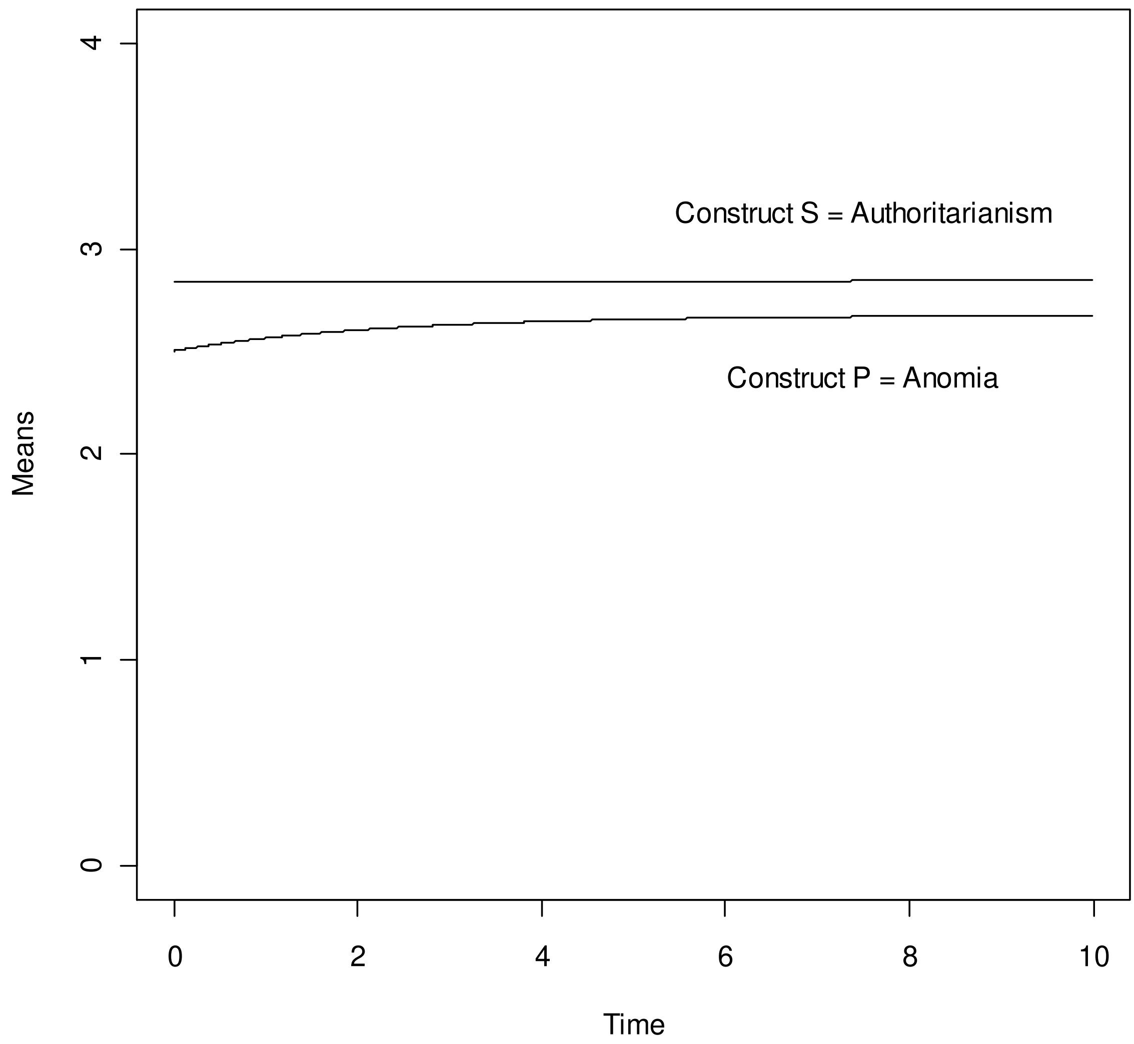
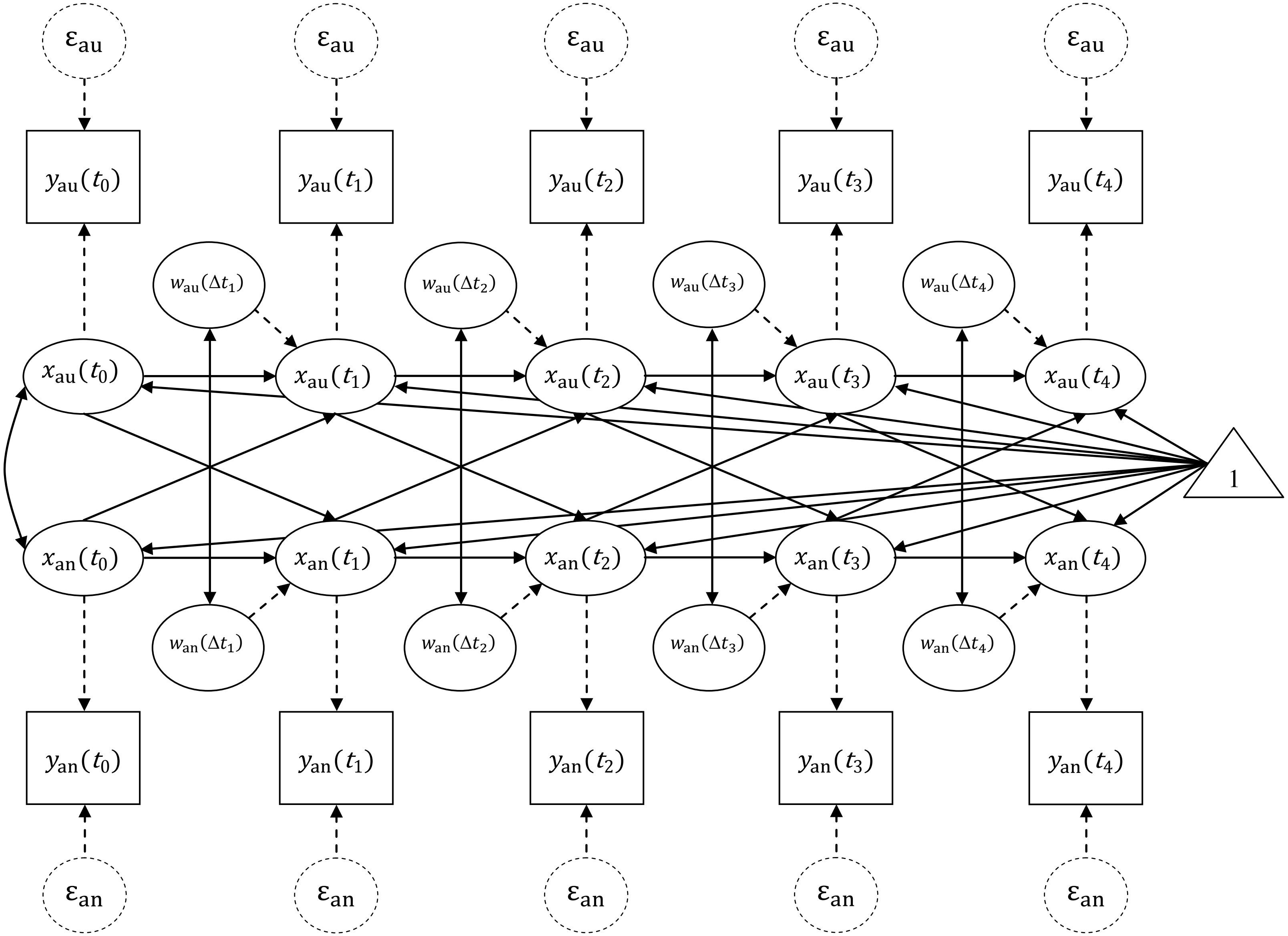

Construct S = Authoritarianism
Construct P = Anomia
Means
Time
0
1
2
3
4
0
2
4
6
8
10

Figure 3



ε_{au}
ε_{au}
ε_{au}
ε_{au}
ε_{au}
$y_{au}(t_0)$
$y_{au}(t_1)$
$y_{au}(t_2)$
$y_{au}(t_3)$
$y_{au}(t_4)$
$w_{au}(\Delta t_1)$
$w_{au}(\Delta t_2)$
$w_{au}(\Delta t_3)$
$w_{au}(\Delta t_4)$
$x_{au}(t_0)$
$x_{au}(t_1)$
$x_{au}(t_2)$
$x_{au}(t_3)$
$x_{au}(t_4)$
1
$x_{an}(t_0)$
$x_{an}(t_1)$
$x_{an}(t_2)$
$x_{an}(t_3)$
$x_{an}(t_4)$
$w_{an}(\Delta t_1)$
$w_{an}(\Delta t_2)$
$w_{an}(\Delta t_3)$
$w_{an}(\Delta t_4)$
$y_{an}(t_0)$
$y_{an}(t_1)$
$y_{an}(t_2)$
$y_{an}(t_3)$
$y_{an}(t_4)$
ε_{an}
ε_{an}
ε_{an}
ε_{an}
ε_{an}

Appendix A

Proof that the unique solution of Equation 5 [$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A}\,\mathbf{x}(t)$] for initial value $\mathbf{x}(t_0) = \mathbf{x}_0$ over any time interval ($\Delta t = t - t_0$) between $\mathbf{x}(t_0)$ and $\mathbf{x}(t)$ is given by Equation 6 [$\mathbf{x}(t) = \mathrm{e}^{\mathbf{A}\cdot(t-t_0)}\mathbf{x}(t_0)$].

The proof consists of first showing that the derivative of Equation 6 is Equation 5 and, second, that taking the initial value in Equation 6 gives $\mathbf{x}(t_0) = \mathbf{x}_0$.

1. Taking the derivative of Equation 6 gives

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{A}\ \mathrm{e}^{\mathbf{A}\cdot(t-t_0)}\mathbf{x}(t_0) = \mathbf{A}\ \mathbf{x}(t).$$

The first rewrite in this equation follows because by definition

$$\mathrm{e}^{\mathbf{A}\cdot t} = \sum_{k=0}^{\infty} \frac{(\mathbf{A}t)^k}{k!} = \mathbf{I} + \mathbf{A}t + \frac{1}{2!}(\mathbf{A}t)^2 + \frac{1}{3!}(\mathbf{A}t)^3 + \cdots$$

and therefore

$$\frac{\mathrm{d}\mathrm{e}^{\mathbf{A}\cdot t}}{\mathrm{d}t} = \mathbf{0} + \mathbf{A} + \mathbf{A}^2 t + \frac{1}{2!}\mathbf{A}^3 t^2 + \frac{1}{3!}\mathbf{A}^4 t^3 + \cdots = \mathbf{A}(\mathbf{I} + \mathbf{A}t + \frac{1}{2!}(\mathbf{A}t)^2 + \frac{1}{3!}(\mathbf{A}t)^3 + \cdots).$$

The second rewrite applies Equation 6.

2. For $t = t_0$, Equation 6 gives

$$\mathrm{e}^{\mathbf{A}\cdot 0}\mathbf{x}(t_0) = \mathbf{x}(t_0) = \ \mathbf{x}_0.$$

Appendix B

Setting Equation 6 [$\mathbf{x}(t) = \mathrm{e}^{\mathbf{A}\cdot(t-t_0)}\mathbf{x}(t_0)$] equal to Equation 2 [$\mathbf{x}(t_i) = \mathbf{A}(\Delta t_i)\mathbf{x}(t_i - \Delta t_i)$] for any arbitrary time point $\mathbf{x}(t) = \mathbf{x}(t_i)$, initial time point $\mathbf{x}(t_i - \Delta t_i) = \mathbf{x}(t_0)$, and time interval $(t - t_0) = \Delta t_i$, we obtain

$$\mathrm{e}^{\mathbf{A}\cdot\Delta t_i}\mathbf{x}(t_0) \ = \mathbf{A}(\Delta t_i)\ \mathbf{x}(t_0),$$

eventually simplifying to Equation 7:

$$\mathbf{A}(\Delta t_i) \ = \mathrm{e}^{\mathbf{A}\cdot\Delta t_i}.$$

Appendix C

Taking the integral of the error part in Equation 10 [$\mathbf{G}\frac{\mathrm{d}\mathbf{W}(t)}{\mathrm{d}t}$] yields

$$\int_{t_0}^{t}\mathbf{G}\frac{\mathrm{d}\mathbf{W}(s)}{\mathrm{d}s}\mathrm{d}s=\int_{t_0}^{t}\mathbf{G}\mathrm{d}\mathbf{W}(s)\ .$$

The variable of integration *s* is only symbolic and replaces *t* in order not to confuse it with the upper limit of integration. Unfortunately, the integral cannot be defined as an ordinary Riemann integral, but can be solved as a Wiener stochastic integral or, more generally, as an Itô stochastic integral with its own rules of integration (e.g., see Arnold, 1974, p. 128–134). Note that $\mathbf{G}$ is independent of time in the differential equation (Equation 10) but its effect varies over the interval. In particular, it is simply multiplied by the matrix exponential derived earlier. Thus we write

$$\int_{t_0}^{t}\mathrm{e}^{\mathbf{A}\cdot(t-s)}\mathbf{G}\mathrm{d}\mathbf{W}(s)\ .$$

If the drift matrix is zero ($\mathbf{A}=\mathbf{0}$), this reduces to $\int_{t_0}^{t}\mathbf{G}\mathrm{d}\mathbf{W}(s)$.

As it is usually the case (e.g., in any structural equation model), the covariance matrix of the error terms corresponds to the expected value of the outer product of the error vectors. In our case this would correspond to

$$cov\left[\int_{t_0}^{t}\mathrm{e}^{\mathbf{A}\cdot(t-s)}\mathbf{G}\mathrm{d}\mathbf{W}(s)\right]=E\left[\left(\int_{t_0}^{t}\mathrm{e}^{\mathbf{A}\cdot(t-s)}\mathbf{G}\mathrm{d}\mathbf{W}(s)\right)\cdot\left(\int_{t_0}^{t}\mathrm{e}^{\mathbf{A}\cdot(t-s)}\mathbf{G}\mathrm{d}\mathbf{W}(s)\right)'\right].$$

The expectation of the product of the two integral forms can be written as one integral form, so that the error covariance matrix corresponds to:

$$\int_{t_0}^{t}\mathrm{e}^{\mathbf{A}\cdot(t-s)}\mathbf{Q}\mathrm{e}^{\mathbf{A}'\cdot(t-s)}\mathrm{d}s=irow\left\{\mathbf{A}_{\#}^{-1}\left[\mathrm{e}^{\mathbf{A}_{\#}\cdot(t-t_0)}-\mathbf{I}\right]row\mathbf{Q}\right\}$$

$$\text{for }\mathbf{Q}=\mathbf{G}\mathbf{G}'\text{ and }\mathbf{A}_{\#}=\mathbf{A}\otimes\mathbf{I}+\mathbf{I}\otimes\mathbf{A}.$$

Mathematical details can be found in Arnold (1974, p. 66–67) and Singer (1990).